

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2002, 16:49:54 ; Search time 2229.59 Seconds
(without alignments)
16519.048 Million cell updates/sec

Title: US-09-786-835-1
Perfect score: 1760
Sequence: 1 ctccagatgtgtagaaga.....gtccatcatcatcagctcca 1760

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl1:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

C	1	76.4	4.3	3203	8	MZETRNBS1	M25397 Zea mays tr
C	2	76.4	4.3	3203	8	ZMB51RPT	X16080 Maize trans
C	3	75.6	4.3	488	11	G71285	C71285 VF044831FM
C	4	62.2	3.5	7218	6	I66494	I66494 Sequence 14
C	5	61	3.5	4648	8	AF326577	AF326577 Zea mays
C	6	53.4	3.0	78101	8	AF031569	AF031569 Zea mays
C	7	53.4	3.0	346296	8	AF090447	AF090447 Zea mays
C	8	53.2	3.0	147198	8	AF466203	AF466203 Zea mays
C	9	51.2	2.9	172805	2	AC066608	AC066608 Homo sapi
C	10	50	2.8	70282	2	AC091141	AC091141 Homo sapi
C	11	49.8	2.8	78167	9	AC092028	AC092028 Homo sapi
C	12	48.8	2.8	5026	6	AX251272	AX251272 Sequence
C	13	48.4	2.8	7503	6	AX346677	AX346677 Sequence
C	14	47.8	2.7	189790	2	AC107420	AC107420 Homo sapi
C	15	47.2	2.7	5416	6	AX346698	AX346698 Sequence
C	16	47	2.7	7753	8	ZMAYSP6	Z11879 Z.mays P ge
C	17	47	2.7	173655	2	OSJN00181	AL662993 Oryza sat
C	18	46.8	2.7	349980	6	AX344566	AX344566 Sequence
C	19	46.8	2.7	349980	6	AX344567	AX344567 Sequence
C	20	46.4	2.6	208396	9	AC003101	AC003101 Homo sapi
C	21	46.2	2.6	28021	3	CELY37E11B	AF125451 Caenorhab
C	22	46.2	2.6	155880	2	AC006739	AC006739 Caenorhab
C	23	46	2.6	349980	6	AX344563	AX344563 Sequence
C	24	46	2.6	349980	6	AX344564	AX344564 Sequence
C	25	45.8	2.6	676	8	PTR286131	AJ286131 Populus t
C	26	45.8	2.6	47577	3	AF396436	AF396436 Tetrahme
C	27	45.8	2.6	173943	2	AC013293	AC013293 Homo sapi
C	28	45.8	2.6	213007	2	AC105762	AC105762 Homo sapi
C	29	45.6	2.6	99003	2	AL390756	AL390756 Homo sapi
C	30	45.6	2.6	191580	2	AC092945	AC092945 Homo sapi
C	31	45.6	2.6	205141	2	AC018679	AC018679 Homo sapi
C	32	45.4	2.6	6412	6	AX251842	AX251842 Sequence
C	33	45.4	2.6	6412	6	AX344228	AX344228 Sequence
C	34	45.4	2.6	6412	6	AX344644	AX344644 Sequence
C	35	45.4	2.6	6412	6	AX348967	AX348967 Sequence
C	36	45.4	2.6	178087	9	AC005089	AC005089 Homo sapi
C	37	45.2	2.6	1200	8	MISCA	V00705 Yeast mitoc
C	38	45.2	2.6	3542	6	YSCMTCG16	L36900 Saccharomyc
C	39	45.2	2.6	14924	6	AX281279	AX281279 Sequence
C	40	45.2	2.6	14924	6	AX345126	AX345126 Sequence
C	41	45.2	2.6	85779	8	SCED11856	AJ011856 Saccharom
C	42	45	2.6	170991	2	AL645585	AL645585 Homo sapi
C	43	45	2.6	199730	2	AC083940	AC083940 Homo sapi
C	44	44.8	2.5	127197	6	AX196294	AX196294 Sequence
C	45	44.8	2.5	146957	2	AC078802	AC078802 Homo sapi

ALIGNMENTS

RESULT 1
MZETRNBS1/c 3203 bp DNA linear PLN 11-MAY-1995
LOCUS MZETRNBS1 Zea mays transposon Bsl.
DEFINITION M25397
ACCESSION M25397.1 GI:168648
VERSION
KEYWORDS insertion sequence; transposon.
SOURCE Zea mays (clone: PK18) 4-week-old seedlings aerial part DNA; and
Ze mays (clone: pui1) 4-week-old seedling aerial part DNA.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3203)
AUTHORS Jin,Y.K. and Bennetzen,J.L.
TITLE Structure and coding properties of Bsl, a maize retrovirus-like
transposon
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86 (16), 6235-6239 (1989)
MEDLINE 89345638
COMMENT Draft entry and computer-readable sequence for [1] kindly provided
by J.L.Bennetzen, 07-JUN-1989.
FEATURES Location/Qualifiers
source 1..3203

/organism="Zea mays"
/db_xref="taxon:4577"
/clone="PK18"
/clone="PJ1"
/haplotype="adh1-S5446"
/tissue_type="seedling aerial part"
/tissue_type="seedlings aerial part"
/dev_stage="4-week-old"
1..302
/note="5' LTR"
305..329
/bound_molety="tRNA primer"
335..2557
/note="ORF1: putative"
/codon_start=1
/product="unknown protein"
/protein_id="AA66269.1"
/db_xref="GI:806301"
/translation="MEPTLOSAMEEQKILREISDRLAQEARMRSESTVTOHSRST
HDLVAMASAPSATLRAELDAPAVATYERLDATPAASARSALESTTATFMDLFDNS
DIYERFNAEVVADDMGGLFGOHAPRLDGGARLLPSPALPASATDTAFSSIVGEC
GTHTYTAALTRPALHKLPLVNAHAKSTPGHNHVCATOTGROARROCRRLVHVR
RLHPTGQLHVRHGRROHVVHQRVLPAGELDGRLRLPMVPAHGRPRRLHOL
HHIDPDGGGGGGHGTTPPERMLSRAGVNAAPNEVDVLRMTLKVLPRTVYCIAGW
IDLDGVAEGADHAAVRFRVDDGCVVEGVCVCSGSWTEIKGFRLKRAVMEVY
VOGAVAGIDVKTDPQVATNVIONLAYVDEFTKHFMAVEKEFERMYHVEDAKVED
KMLTKGEAPSVQGVYVVAHNLVQALFMDGRASDAYVLEEMONGPEPPEVFTY
TLKYLKGNMAEESTILVPGDITGVKLGDIISADTRLBEDPLKIDOSALTGFCIC
SIYAGMLVEFTVMPIDOMVYRPRIKLLVLLIGGIFLAMPVTVSVTMSIGATYLAQO
GATIRMTTIEEMAGMDVPCSDKTGTLPMTKLIVIKSLVDVFOGADODAVLMDARA
SCTRKNDAIEATIVSMLAAPKEACAGVOEIOFLFPNPNDRKRTAVTYSLLIYALSPGA
"
660..1088
/note="ORF3: putative"
/codon_start=1
/product="unknown protein"
/protein_id="AA66270.1"
/db_xref="GI:806302"
/translation="MLRMLPTIGVCSDSMLTLTLEHVFYFLPPLPSPHPRILPS
PASTASVARSSTPRLLSPVLIHFTNPMNMTPEVROATTECALKFNAROLDGLDAN
VDYACMFADCTSLGISTCVSMVYVNTSTFNMSYAPN"
2608..3060
/note="ORF2: putative"
/codon_start=1
/product="unknown protein"
/protein_id="AA66271.1"
/db_xref="GI:806303"
/translation="MYNALINRCHTDGNNMEKAYIEMKKRIIPDNVTCNTLMRESC
LIGQLDEARLILIDMTKRSTAIISHDGNFSTVWDPGHKNSSLGITSCILWGGCQILY
RMVAGHDGPGILHYCSILIOBEKNTHELHGRQETIIPSSISIVFLGSC"
2902..3203
/note="3' LTR"
BASE COUNT 738 a 813 c 870 g 782 t
ORIGIN
Query Match 4.3%; Score 76.4; DB 8; Length 3203;
Best Local Similarity 69.3%; Pred. No. 4.9e-09;
Matches 104; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

RESULT 2
ZMBISRET/c 3203 bp DNA linear PLN 12-SEP-1993
LOCUS Maize transposable element Bst1.
DEFINITION
ACCESSION X16080.1 GI:22198
VERSION X16080.1
KEYWORDS long terminal repeat; retrotransposon; reverse transcriptase;
transposable element.
SOURCE
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 3203)
AUTHORS Johns,M.A., Babcock,M.S., Fuerstenberg,S.M., Fuerstenberg,S.I.,
Freeeling,M. and Simpson,R.B.
TITLE An unusually compact retrotransposon in maize
JOURNAL Plant Mol. Biol. 12, 633-642 (1989)
FEATURES
source
location/Qualifiers
1..3203
/organism="Zea mays"
/db_xref="taxon:4577"
/dev_stage="seedlings"
1..302
/note="5'-long terminal repeat"
247..254
/note="pot. TATA box"
275..280
/note="pot. polyadenylation signal"
305..315
/note="minus strand primer binding site"
335..2557
/note="ORF 1 polyprotein (AA 1-740)"
/codon_start=1
/protein_id="CMA34210.1"
/db_xref="GI:22199"
/db_xref="WazedB:Variation/15764"
/db_xref="Swiss-Prot:P15718"
/translation="MEPTLOSAMEEQKILREISDRLAQEARMRSESTVTOHSRST
HDLVAMASAPSATLRAELDAPAVATYERLDATPAASARSALESTTATFMDLFDNS
DIYERFNAEVVADDMGGLFGOHAPRLDGGARLLPSPALPASATDTAFSSIVGEC
GTHTYTAALTRPALHKLPLVNAHAKSTPGHNHVCATOTGROARROCRRLVHVR
RLHPTGQLHVRHGRROHVVHQRVLPAGELDGRLRLPMVPAHGRPRRLHOL
HHIDPDGGGGGGHGTTPPERMLSRAGVNAAPNEVDVLRMTLKVLPRTVYCIAGW
IDLDGVAEGADHAAVRFRVDDGCVVEGVCVCSGSWTEIKGFRLKRAVMEVY
VOGAVAGIDVKTDPQVATNVIONLAYVDEFTKHFMAVEKEFERMYHVEDAKVED
KMLTKGEAPSVQGVYVVAHNLVQALFMDGRASDAYVLEEMONGPEPPEVFTY
TLKYLKGNMAEESTILVPGDITGVKLGDIISADTRLBEDPLKIDOSALTGFCIC
SIYAGMLVEFTVMPIDOMVYRPRIKLLVLLIGGIFLAMPVTVSVTMSIGATYLAQO
GATIRMTTIEEMAGMDVPCSDKTGTLPMTKLIVIKSLVDVFOGADODAVLMDARA
SCTRKNDAIEATIVSMLAAPKEACAGVOEIOFLFPNPNDRKRTAVTYSLLIYALSPGA
"
2497..3057
/note="ORF 2"
2889..2901
/note="plus strand primer binding site"
misc_feature
2902..3203
/note="3'-long terminal repeat"
3148..3155
/note="pot. TATA box"
3176..3181
/note="pot. polyadenylation site"
BASE COUNT 738 a 813 c 870 g 782 t
ORIGIN
Query Match 4.3%; Score 76.4; DB 8; Length 3203;
Best Local Similarity 69.3%; Pred. No. 4.9e-09;
Matches 104; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

KEYWORDS

KEYWORDS

VERSION

SOURCE
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 4648)
REFERENCE
AUTHORS Selinger,D.A. and Chandler,V.L.
TITLE B-Bolivia, an Allele of the Zea mays b1 Gene with Variable Expression, Contains a High Copy Retrotransposon-Related Sequence Immediately Upstream
JOURNAL Plant Physiol. 125 (3), 1363-1379 (2001)
PUBMED 11244116
REFERENCE
2 (bases 1 to 4648)
AUTHORS Selinger,D.A. and Chandler,V.L.
TITLE Direct Submission
JOURNAL Submitted (06-DEC-2000) Department of Plant Sciences, University of Arizona, 303 Forbes, Bldg. 36, Tucson, AZ 85721, USA
LOCATION/Qualifiers
1..4648
FEATURES
source
1..4648
/organism="Zea mays"
/db_xref="taxon:4577"
complement(<1..>858)
/note="start site undetermined"
/codon_start=1
/evidence=not_experimental
/product="gag protein"
/protein_id="AAK16154.1"
/db_xref="GI:13236411"
/translation="GPRRLNRRTAAARPAPOEPSPDEARDTGGAPTCFRISQSA
NDRRGAESPSPARSDHAGQGHGLQDEPPLQOROPEDRTGGLGRQGPLOHRA
QENPRPQHDYAPQHGAPFGQINLYDERSPLPHLQVTPWPNFAGAPKNGSTD
PAQYMSGYAVASAGDEATMAKSFILAEGLALMTFTLPLSIDSMRLDKFLL
NMGYRPPDIALAELSLCKLEKLETLREYRKRLTLKSQLPVDDQIAIHVYANASGR
RPLQPIYQGS"
complement(<1..973)
/organism="Zea mays"
/db_xref="taxon:4577"
/transposon="Shadowsawn1 retrotransposon"
1610..2103
/note="similar to the B-peru aleurone specific promoter
region: aleurone specific expression"
/evidence=experimental
2104..2111
TATA_signal
/evidence=experimental
join(2142..2166,2392..2672,3198..3452,3533..3629,
3790..3804,3939..3995,4643..>4648)
/product="anthocyanin regulatory B protein"
join(2530..2672,3198..3452,3533..3629,3790..3804,
3939..3995,4643..>4648)
/note="B-Bolivia allele"
/codon_start=1
/evidence=not_experimental
/product="anthocyanin regulatory B protein"
/protein_id="AAK16153.1"
/db_xref="GI:13236410"
/translation="MALSAAPQAEELQAPGRPLRKOLAARSINMVALFWSISST
ORRVLMTDRFNGEYVTRKISHVELTADOLMRELYEALQSGECRRRA
RPGVSLSPEDLGTMEYVYVCMYAFPLGGGLPGRSSASNEHWLCAHLAGSKDFR
ALAKSKCITIVICPLMGVLELGTDKYP"
BASE COUNT 988 a 1153 c 1284 g 1223 t
ORIGIN
Query Match 3.5%: Score 61; DB 8; Length 4648;
Best Local Similarity 66.3%: Pred. No. 6e-05;
Matches 124; Conservative 0; Mismatches 50; Indels 13; Gaps 2;
Oy 846 tagtgcctaataatagccaagaagtttagaaatgatacaactaatgttagctataccctt 905
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 4461 TAGCAGTTAAATTAAGCTAAGAGTTTAGAACAATCAAGTTAATATAGTTACTTCT 4520
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Oy 906 taataagatataatgttagat-----ggccaacctagctaaaa-----ccagctt 952
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||

Db 4521 AAAATAGTATTAGTTGTAGCTACTAAGCTAAACCAGCTAACACTTATTAGCAACT 4580
Oy 953 aacagtagctctagatagtttagagacatgacctaaacacatatatagagccctaccgc 1012
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 4581 AACATTAGCTTAGAGAGTTTGAAATAGGCTGTACACCACTACTGTAAGGCAATTACC 4640
Oy 1013 cggatcc 1019
|| ||
Db 4641 AGTGCC 4647
RESULT 6
AF031569
LOCUS 78101 bp DNA linear PLN 05-FEB-1998
DEFINITION Zea mays 22-kDa alpha zein gene cluster, complete sequence.
ACCESSION AF031569
VERSION AF031569.1 GI:2832242
KEYWORDS
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 78101)
AUTHORS Liaca,V. and Messing,J.
TITLE Structure and organization of the 22-kDa alpha zein gene cluster in Zea mays
JOURNAL Unpublished
2 (bases 1 to 78101)
AUTHORS Liaca,V. and Messing,J.
TITLE Direct Submission
JOURNAL Submitted (27-OCT-1997) Waksman Institute, Rutgers, The State University of New Jersey, P.O. Box 759, Piscataway, NJ 08855-0759, USA
FEATURES
source
1..78101
/organism="Zea mays"
/db_xref="taxon:4577"
/chromosome="4"
/map="4s"
/note="cosmid contig; inbred line BSSS53"
1..39083
/organism="Zea mays"
/db_xref="taxon:4577"
/chromosome="4"
/map="4s"
/clone="III 4H10"
/note="inbred line BSSS53"
1..541
/note="similar to Zea mays c1 locus myb homolog CDS and gp4 promoter region"
1471..2270
/gene="azs22-1"
/note="22-kDa alpha zein 1"
/pseudo
5582..6382
/gene="azs22-2"
/note="22-kDa alpha zein 2"
/pseudo
9689..9766
/gene="azs22-3"
/note="22-kDa alpha zein 3; truncated"
/pseudo
10702..10770
/note="microsatellite"
/rpt_type=tandem
/rpt_unit=AT
15639..15694
/note="microsatellite"
/rpt_type=tandem
/rpt_unit=AT
18119..18919
/gene="azs22-4"

SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
TITLE 1 (bases 1 to 346296)
1 (bases 1 to 346296)
Song, R., Liaca, V., Linton, E. and Messing, J.
Genomic imprinting as a Rescue for a Compact Gene Cluster of Zea
mays L.
JOURNAL Unpublished
REFERENCE 2 (bases 80292 to 116863)
AUTHORS Song, R., Liaca, V. and Messing, J.
JOURNAL Direct Submission
TITLE Submitted (04-SEP-1998) Waksman Institute, Rutgers University, 190
3 (bases 1 to 65155)
Liaca, V., Lou, A., Young, S. and Messing, J.
JOURNAL Direct Submission
TITLE Submitted (10-NOV-1998) Waksman Institute, Rutgers University, 190
4 (bases 1 to 346296)
Song, R., Liaca, V. and Messing, J.
JOURNAL Direct Submission
TITLE Submitted (12-APR-2001) Waksman Institute, Rutgers University, 190
Frelinghuysen Road, Piscataway, NJ 08854-8020, USA
Sequence update by submitter
5 (bases 1 to 346296)
Song, R., Liaca, V. and Messing, J.
JOURNAL Direct Submission
TITLE Submitted (24-APR-2001) Waksman Institute, Rutgers University, 190
Frelinghuysen Road, Piscataway, NJ 08854-8020, USA
Amino acid sequence updated by submitter
On Apr 12, 2001 this sequence version replaced gi:4416300
gi:4140643.
FEATURES
source Location/Qualifiers
1..346296
/organism="Zea mays"
/db_xref="taxon:4577"
/chromosome="4"
/map="4s"
/clone="cosmid IT.2E10"
/note="inbred line BSS53"
1..65155
/organism="Zea mays"
/db_xref="taxon:4577"
/chromosome="4"
/map="4s"
/clone="cosmid III.3C12; V.9D7"
/note="inbred line BSS53"
join(<1256..1397,2128..>2300)
/product="hypothetical protein"
join(1256..1397,2128..2300)
/codon_start=1
/product="hypothetical protein"
/protein_id="AAD20309.1"
/db_xref="GI:4416304"
/translation="MYKILGVLEHDDTDADRSSEKGLDVGCTETTEPESTFGVAYW
KGGSATVPSNAQKVFYTWRMPEPFLSKLLPLHSHIDQNTGTSWTRFVNDGGLIRLQ
IP"
1261..185036
/organism="Zea mays"
/db_xref="taxon:4577"
/chromosome="4"
/map="4s"
/clone="BAC 204"
/note="inbred line BSS53"
join(<3179..4286,5020..5155,5266..5443,5892..>6020)
/product="hypothetical protein"
join(3179..4286,5020..5155,5266..5443,5892..6020)
/codon_start=1
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="AAD20310.1"

/db_xref="GI:4416305"
/translation="MVCKGTVFSGSDTSVLETFQTVGAGSVASNLACQCTCSAVFGGD
KEMANCKVEHASSGSRTRVVASGNGHTERTGCGGCGSNGCPMAAEDPRAKNGKSGG
CGSCGCGCGGAGCGCGTMANNGSRSGGGLTLLSNMAGGQAOHSAHGCGGCGGC
GSGVITBESKANNMAMSGCGCGCGCGGCGSSMLITVSKANNMMSGCGCGCGGCGSGM
VNEGSKANNAMSGCGCGCGCGGCGGLFNASTAAGGLINKSAGCGGCGCGCGGCGG
CGSGMAIEGFANNAKSGCGCGCGCGCGGCGGLTFSSAAGAQDSRSGCGCGCGGCGG
GSMAVEGNGGSHAKSGCGCGCGCGCGGCGGATTTLSLMPADVYTAEGESGTA
RCEELVGPAPPTPAFTTFYGGRAKSLQRFACAGCDGCGDELSPFLMPVIMEVLIN
VYADPSRLHAKDLVNPQHDFVETCCCSSTPMLAAASLQNAFEGTKEQPAATQ
TCCCCFOK"
join(<11112..11154,11262..11396,12254..12344,14272..14410,
14838..14890,18185..18230,18589..>18657)
/product="hypothetical protein"
join(11112..11154,11262..11396,12254..12344,14272..14410,
14838..14890,18185..18230,18589..18657)
/codon_start=1
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="AAD20311.1"
/db_xref="GI:4416306"
/translation="MDVGMRPMTATRVDERESRAHGDGLTGMSSVATGEAETIHGGC
EEVEGQHDMAPOREVILGRMDIGSTSRCSVGVSVAPPRGLARYTEPELLEMLSF
DHRCPYAGLGVEITFEKRYTLDPSLTHYVESAVTNLYELNDRDGLTSQLLRSI
GMEDPKVTVLKNQEAISLMSISKDQAAAK"
15142..17012
/note="similar to cell division protein FTS2"
22068..22142
/rpt_type=dispersed
24303..24307
/note="sequence target duplication"
24306..31758
/rpt_family="zeon-1 retroelement"
/rpt_type=dispersed
24308..24976
/product="gag protein"
25532..26569
/codon_start=1
/product="gag protein"
/protein_id="AAD20308.1"
/db_xref="GI:4416303"
/translation="MSYQATVSSYGNTFVMAKSPVMVRSVAQOTWYSLRPGTITSM
OKLMDLITSEFGFOTKPYTAQALROCTDDHEEYLDQAVRRRLRLRAQAPVYNEIYI
EAMIKGLRPGSAQYFAARKPQTLLELLQKMEYIRADNDFQRREARFSEMRGF
GGRFYRHAHRSYHNSTONDDKSSQOQSSQASQOQSSRPAPRGRGARFGER
FGDOPRRICLFCGKNGHTTRMCVHTTIOKEIEAQAQAPKOIMHTASHSPIY
PEYGNHPAVSVASASQQAASHQPPPPPLQGGQPEGGQYQAQHORDFRQSEARTV
NSTVPSKHIY"
31090..31758
31759..31763
/note="sequence target duplication"
32800..32855
/rpt_type=dispersed
36025..37602
/note="pnp(p10)20725 RFLP marker"
join(<36233..36579,37382..>37925)
/product="hypothetical protein"
join(36233..36579,37382..37925)
/codon_start=1
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="AAD20312.1"
/db_xref="GI:4416307"
/translation="MGSPFGHVLPGTLLFLAVGLMRVSSAARHAAPSSFRVRAWSPA
PAPGPRLLELYVAGGALADMCVEVLYSTHLRIFAGGEVNPANHLNLDLEHGMILMEF
LRGALALASQIMRPHRLPDGALCIYAARFAFELVLFPHSTTHMGLBEYVYLLVL
LVGLCYAAVATGLILPDSFPADLASGVIVAQGLMYQNALALYGMPLPGCARDDA
HVECRSRAQDERAEQALISQLEFALVLLAFILALGYCAVAARARGREPPELLPAVHRRRA
AVELRCROGDGDAAMECAI"
complement(38797..38910)
/rpt_type=dispersed
misc_feature complement(39147..39151)

repeat_region	/note="target sequence duplication"
LTR	/rpt_family="gypsy/Ty3-type retrotransposon"
LTR	/rpt_type=dispersed
CDS	complement(39152..49223)
	complement(39152..42406)
	complement(42452..45283)
	/note="gypsy-type polypeptide"
	/codon_start=1
	/pseudo
	complement(45973..49223)
	complement(49224..49228)
	/note="target sequence duplication"
	complement(50243..50571)
	/note="similar to gypsy 1"
	complement(50572..60252)
	/note="nested in grande 1 retroelement"
	/rpt_family="copia-type retroelement PREM-2"
	/rpt_type=dispersed
	complement(50572..50576)
	/note="target sequence duplication"
	complement(50577..52000)
	/note="copia-type retroelement PREM-2"
	complement(52949..56140)
	/product="copia-type pol polypeptide"
	complement(52949..56140)
	/codon_start=1
	/product="copia-type pol polypeptide"
	/protein_id="AAD20307.1"
	/db_xref="gi:4416302"
	/translation="MTGKRMSSSEYKNDOPRATFEGNGOVLKGLTASPDH ISNELVSLDLYNLISVQLCOMGNCFLPTDGVPRSPDSIAFKVLSOLYLV FDRRLDLAKTNGMGLMRRLAHVSMKMLHKLKGEHLIGTNVFEEDRCSA OAGVOYTHHNRKMLMTDRLELHMLDFGLATISIGSKYCLIVDDYSRFTVA FLOESQTOETLKGFLRAONEFGRLRKISDSGTERNSOIEFLEDEGIEKFS PYTPQONGVERKNTLDMARMTDEKTPDREAEVAENLACYAENLYLRLIKK SYELTGKRPNIISFYFGSKFLIKGRSKSPAPKEGFLIGDSNTRAYRENN STGLSEADVDFEPTNGSOEYOVDLIGEOAPALRMMSIGDNCPRKSEEPST ODOSSMSQSPPTONEDEAONDEONEDERPODSDSGDGTNDOKEDDEEPRP HPRHQAORHPVDTLIGLHKGKTSRAHCEHISFVSSIEPHVEEALDSDSD VVAQOELNFTNEVHLVPRPNONVGTWVPRNKODEGVYTRNKARLVAGYS VEGIDFETVAPVAREILILVATYAHGKFLVOMDYKSAFLMPLKEEYVDEP FEDEEYPMHVRILSKALYGLKQAPRAWECGRDFLIANGFVGADPTLPKTLN FVCOIYVDIDIFGTSKSTCEFSRIMQREKMSMGILKYFLGVYKLOEFGTFC TKYODLITKFGMKDAPKIKYPMGTNGHLDIDTGSKSDVQVYMSISLYLCA DIMLSCARFQSDPKRESHTTAYKRLIRLYLATYPRKELMTPRSTEDLIGYSD GCKINRSTGTCQFLGRSLVSMASKONSVALSTAEVYIAAGHCCQOLIMRQTL DYGLKLRVPLDCNESAIKMADNPVESHRSKHAIRYHFLRDHQKGDIEISYIN DQLDIFTRPLDEOSFRLRHELNILDSRNFEC" complement(<56208..>58886)

FEATURES	source
LOCUS	AF466203
DEFINITION	Zea mays clone ZM8BC_0092E12 putative RIR2 orf3, putative gypsy-type reticulated RIR2, putative orf1, regulatory protein, putative pol protein, putative pol protein, putative gag protein, putative pol protein, putative cis-zeatin O-glucosyltransferase, putative gag-pol precursor -orf1, putative pipol, putative gag-pol precursor -orf2, and putative pipol genes, complete cds.
ACCESSION	AF466203.1
VERSION	GI:18568234
KEYWORDS	
SOURCE	Zea mays.
ORGANISM	Zea mays.
REFERENCE	Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
AUTHORS	1 (bases 1 to 147198)
JOURNAL	Ramakrishna, W., Sam Miguel, P., Emberton, J. and Bennetzen, J. Direct Submission
REFERENCE	Submitted (07-JUN-2002) Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA
AUTHORS	2 (bases 1 to 147198)
TITLE	Liaca, V., Linton, E.W., Young, S., Kovchok, S. and Messing, J. Direct Submission
JOURNAL	Submitted (07-JAN-2002) Rutgers, The State University of New Jersey, The Plant Genome Initiative at Rutgers - Waksman Institute, 150 Frelinghuysen Road, Piscataway, NJ 08854, USA
FEATURES	Location/Qualifiers
source	1..147198
	/organism="Zea mays"
	/cultivar="B73"
	/db_xref="taxon:4577"
	/chromosome="2"
	/map="2S; 2.03"
	/clone="ZM8BC_0092E12"
	join(5148..5420,5843..5923,6003..6658,6741..7344,7419..7764,7846..>8336)
	/gene="Z092E12.1"
	/product="putative RIR2 orf3"
	<5148..>8336
	/gene="Z092E12.1"
	/note="similar to Zea mays RIR2 orf3 GB:AF391808.2 GI:17082476"
	/evidence=not_experimental
	join(5148..5420,5843..5923,6003..6658,6741..7344,7419..7764,7846..8336)
	/gene="Z092E12.1"
	/codon_start=1
	/evidence=not_experimental
	/product="putative RIR2 orf3"
	/protein_id="AAL75972.1"
	/db_xref="GI:18568235"
	/translation="MGSVAASPPPGVYVSVSHRNGFGVPTFRFMAIRMFHIXGVEIH NLRPNISQAIIFYAVCEGYLTGTAIPMDHTLIFAEFLPTGERKVSAMGKPPSH AFSQLMRPDQGCVTLEVGWHRKPSLPVPPDADGAARYAAEKKKKDAEKARAE RRRARNALEKCRQDREKGLPRPSPETPDDDDDDDDEGDMAARLGLSGICGCE PSSQSPGPIPVGVGAGSGWPEARGRPERSDPSGAGAEVAPGQVRASAPGPTL VPAHGGDPOVLTIVPEPSAPRASKAIVAPKLPVKRSSAAVSGAGIOEFSPOARWIMA BSKRSRGODIAPRKAKITASYACASAPGAYVVOPTLSGVSPQCARASPVSGEVP FASSSDEALIVGEAGANVSSPLVYPMALAPATLEVAIVGERVVAASGVMAVS APSASEVGVAVRSVOPGSDLIWVRSPKARROLRFRTREASDPVFTLDDDEQEDSW DLEKREAEATVGSLSRSLVEFCDKVPKLIQDILDRSAASSFTRCVDVWDSLSRLT SLDGAAARLSQDAAKVAADLOLCADIRAEAAACAGQORSEFYVVERDQSRRA VEESEAEETLAADLAAYOAAVEORARAELELADSEAKALAEALGAEOREADLA AMSEAAVDVYRIIGSDVPSGSSPOSRLQALGPYAKGTIREALHHGVRAAFNTASHY VNDLEVSEGYCLPDEDEAALAEVQRLDVAANAPSAALATTPEAETLTPAPSLAEVD LINDGGCTGCAPSSCGGA"
	repeat_region
	/ipr_unif_cga
	6215..6254

[illegible]

```

repeat_region
repeat-1
  /note="At-rich"
  /evidence=not_experimental
  join(<30970..31268,31897..>32083)
  /gene="Z092E12.6"
  /product="putative pol protein"
  <30970..>32083
  /gene="Z092E12.6"
  /codon_start=1
  /evidence=not_experimental
  /product="putative pol protein"
  /protein_id="AAL75976.1"
  /db_xref="GI:18568239"
  /translation="MTPEELIRKRFVSGCMYKEARVVDALNGPLSYEPDLIALAT
SYEALPESKVAOYEAGLNNENMALIRKTKALRGCEPSKSKTGEALLQMRCK
FYPVDNCKEHLITAKNDENVASLNAQKCKSDPDKFRPADAYTTIGRHSIKDGLG
P"
  complement(42008..42064)
  /rpt_unit=cag
  /evidence=not_experimental
  join(<44689..44868,45357..45548,46078..>46230)
  /gene="Z092E12.8"
  /product="putative pol protein"
  <44689..>46230
  /gene="Z092E12.8"
  /note="similar to maize retrotransposon Opie-2 pol gene
GB:068408.1 GI:1657766"
  /evidence=not_experimental
  join(<44689..44868,45357..45548,46078..46230)
  /gene="Z092E12.8"
  /codon_start=1
  /evidence=not_experimental
  /product="putative pol protein"
  /protein_id="AAL75977.1"
  /db_xref="GI:18568240"
  /translation="MTGEKKMFITYVKNKDSODTIIFGDGNOGVKGLKIAITNHS
TSNVLVESIGYNLFPDKSETOGLKFLFLRAQNEFLKVKIRSDGSKFNLOVEE
YIEPEGVKEHSAAPYTPQNDNVYDRFLKKRVIKGDKKKMLRKHHGSLQKSR
RSNCKITPLTRFWY"
  complement(join(<51501..51584,52822..53786,53976..>54270)))
  /gene="Z092E12.9"
  /product="putative gag protein"
  complement(<51501..>54270)
  /gene="Z092E12.9"
  /note="similar to maize retrotransposon Opie-2 gag gene
GB:068408.1 GI:1657766"

mRNA
  /gene="Z092E12.9"

gene
  /note="similar to maize retrotransposon Opie-2 gag gene
GB:068408.1 GI:1657766"

Query Match      3.0%; Score 53.2; DB 8; Length 147198;
Best Local Similarity 82.4%; Pred. No. 0.011;
Matches 61; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY      856 aattagccaaaagtgatgatactatgtagctatacccttaataagta 915
      ||||||| || ||||||| || ||||||| ||||||| ||||||| ||
Db 24141 AATTAGCTAGAGGTTTGAACCCATCATTAATTGTTAGTTATACCCCTAATAACTA 24200

OY      916 ttaactgtatagctg 929
      || ||||||| ||
Db 24201 CTAGTTGTTAGCTG 24214

RESULT      9

```


LOCUS	AC066608	172805 bp	DNA	linear	PRI 11-FEB-2001
DEFINITION	Homo sapiens chromosome 3 clone RP11-785A7 map 3p, complete sequence.				
ACCESSION	AC066608				
VERSION	AC066608.5	GI:12745084			
KEYWORDS	HTG.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 172805)				
REFERENCE	Xiong,H., Zhou,Y., Dong,H., Liu,W., Chen,B., Zhang,C., Zhang,Y., Cai,Z., Ying,H.F., Wang,H., Gu,W., Zhu,G., Tu,Y., Zhang,X., Jia,J., Shen,H., Zhang,D., Wu,C., Lu,G., Zhong,M., Jiang,H., Ren,S., Fu,G., Chen,Z. and Huang,M.				
TITLE	Chromosome 3p genomic sequence				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 172805)				
AUTHORS	Zhang,X., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B., Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y., Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y., Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D., Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J. and Yang,H.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-APR-2000) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China				
REFERENCE	3 (bases 1 to 172805)				
AUTHORS	Bao,W., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,T., Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D., Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,F., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y., Luo,C., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L., Song,S., Sun,W., Sun,Y., Tao,X., Tao,R., Wang,H., Wang,J., Wang,J., Wang,J., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X., Wang,T., Wu,D., Wu,Q., Xie,F., Xian,Z., Xue,Y., Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N., Yu,J. and Yang,H.				
TITLE	Direct Submission				
JOURNAL	Submitted (11-FEB-2001) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China				
COMMENT	On Feb 11, 2001 this sequence version replaced gi:8101281.				
	-----Genome Center				
	Center:Beijing Center				
	Center code:Beijing				
	Website:http://hgsc.igtp.ac.cn				
	http://www.genomics.org.cn				
	Contact:hgsc@igtp.ac.cn				
	-----Project Information				
	Center project name:18 project				
	Center clone name: RP11-785A7				
	-----Summary Statistics				
	Sequencing vector: pUC18; 100% of reads				
	Chemistry: Dye-terminator; HT 55% of reads				
	Chemistry: Dye-terminator Big Dye; 45% of reads				
	Assembly program: Phrap; version 0.990329				
	Consensus quality: 176043 bases at least Q40				
	Consensus quality: 176221 bases at least Q30				
	Consensus quality: 176261 bases at least Q20				
	Insert size: 172805; sum-of-contigs				
	Quality coverage: 8.12x in Q20 bases;sum-of-contigs				

FEATURES	Location/Qualifiers				
source	1..172805				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="3"				
	/map="3p"				
	/clone="RP11-785A7"				

[illegible]

* NOTE: This record contains 84 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 712: contig of 712 bp in length
* 713 812: gap of 100 bp
* 813 1577: contig of 765 bp in length
* 1578 1677: gap of 100 bp
* 1678 2430: contig of 753 bp in length
* 2431 2530: gap of 100 bp
* 2531 3317: contig of 787 bp in length
* 3318 3417: gap of 100 bp
* 3418 4146: contig of 729 bp in length
* 4147 4246: gap of 100 bp
* 4247 4979: contig of 733 bp in length
* 4980 5079: gap of 100 bp
* 5080 5822: contig of 743 bp in length
* 5823 5922: gap of 100 bp
* 5923 6660: contig of 738 bp in length
* 6661 6760: gap of 100 bp
* 6761 7506: contig of 746 bp in length
* 7507 7606: gap of 100 bp
* 7607 8343: contig of 737 bp in length
* 8344 8443: gap of 100 bp
* 8444 9176: contig of 733 bp in length
* 9177 9276: gap of 100 bp
* 9277 10015: contig of 739 bp in length
* 10016 10115: gap of 100 bp
* 10116 10862: contig of 747 bp in length
* 10863 10962: gap of 100 bp
* 11712 11711: contig of 749 bp in length
* 11812 12514: contig of 703 bp in length
* 12515 12614: gap of 100 bp
* 12615 13370: contig of 756 bp in length
* 13371 13470: gap of 100 bp
* 13471 14196: contig of 726 bp in length
* 14197 14296: gap of 100 bp
* 14297 15033: contig of 737 bp in length
* 15034 15133: gap of 100 bp
* 15134 15920: contig of 787 bp in length
* 15921 16020: gap of 100 bp
* 16021 16737: contig of 717 bp in length
* 16738 16837: gap of 100 bp
* 16838 17569: contig of 732 bp in length
* 17570 17669: gap of 100 bp
* 17670 18421: contig of 752 bp in length
* 18422 18521: gap of 100 bp
* 18522 19258: contig of 737 bp in length
* 19259 19358: gap of 100 bp
* 19359 20138: contig of 780 bp in length
* 20139 20238: gap of 100 bp
* 20239 21031: contig of 793 bp in length
* 21032 21131: gap of 100 bp
* 21132 21849: contig of 718 bp in length
* 21850 21949: gap of 100 bp
* 21950 22706: contig of 757 bp in length
* 22707 22806: gap of 100 bp
* 22807 23565: contig of 759 bp in length
* 23566 23665: gap of 100 bp
* 23666 24380: contig of 715 bp in length
* 24381 24480: gap of 100 bp
* 24481 25218: contig of 738 bp in length
* 25219 25318: gap of 100 bp
* 25319 26012: contig of 694 bp in length

* 26013 26112: gap of 100 bp
* 26113 26842: contig of 730 bp in length
* 26843 26942: gap of 100 bp
* 26943 27673: contig of 731 bp in length
* 27674 27773: gap of 100 bp
* 27774 28499: contig of 726 bp in length
* 28500 28599: gap of 100 bp
* 28600 29372: contig of 773 bp in length
* 29373 29472: gap of 100 bp
* 29473 30239: contig of 767 bp in length
* 30240 30338: gap of 100 bp
* 30340 31077: contig of 738 bp in length
* 31078 31177: gap of 100 bp
* 31178 31896: contig of 719 bp in length
* 31897 31996: gap of 100 bp
* 31997 32782: contig of 786 bp in length
* 32783 32882: gap of 100 bp
* 32883 33617: contig of 735 bp in length
* 33618 33717: gap of 100 bp
* 33718 34429: contig of 712 bp in length
* 34430 34529: gap of 100 bp
* 34530 35242: contig of 713 bp in length
* 35243 35342: gap of 100 bp
* 35343 36085: contig of 743 bp in length
* 36086 36185: gap of 100 bp
* 36186 36899: contig of 714 bp in length
* 36900 36999: gap of 100 bp
* 37000 37734: contig of 735 bp in length
* 37735 37834: gap of 100 bp
* 37835 38594: contig of 760 bp in length
* 38595 38694: gap of 100 bp
* 38695 39445: contig of 751 bp in length
* 39446 39545: gap of 100 bp
* 39546 40277: contig of 732 bp in length
* 40278 40377: gap of 100 bp
* 40378 41118: contig of 741 bp in length
* 41119 41218: gap of 100 bp
* 41219 41946: contig of 728 bp in length
* 41947 42046: gap of 100 bp
* 42047 42768: contig of 722 bp in length
* 42769 42868: gap of 100 bp
* 42869 43601: contig of 733 bp in length
* 43602 43701: gap of 100 bp
* 43702 44446: contig of 745 bp in length
* 44447 44546: gap of 100 bp
* 44547 45197: contig of 651 bp in length
* 45198 45297: gap of 100 bp
* 45298 46021: contig of 724 bp in length
* 46022 46121: gap of 100 bp
* 46122 46861: contig of 740 bp in length
* 46862 46961: gap of 100 bp
* 46962 47706: contig of 745 bp in length
* 47707 47806: gap of 100 bp
* 47807 48644: contig of 838 bp in length
* 48645 48744: gap of 100 bp
* 48745 49491: contig of 747 bp in length
* 49492 49591: gap of 100 bp
* 49592 50332: contig of 741 bp in length
* 50333 50432: gap of 100 bp
* 50433 51170: contig of 738 bp in length
* 51171 51270: gap of 100 bp
* 51271 51999: contig of 729 bp in length
* 52000 52099: gap of 100 bp
* 52100 52830: contig of 731 bp in length
* 52831 52930: gap of 100 bp
* 52931 53640: contig of 710 bp in length
* 53641 53740: gap of 100 bp
* 53741 54467: contig of 727 bp in length
* 54468 54567: gap of 100 bp
* 54568 55301: contig of 734 bp in length
* 55302 55401: gap of 100 bp
* 55402 56150: contig of 749 bp in length
* 56151 56250: gap of 100 bp

```

*      56251      57001: contig of 751 bp in length
*      57002      57101: gap of 100 bp
*      57102      57819: contig of 718 bp in length
*      57820      57919: gap of 100 bp

Query Match      2.8%; Score 50; DB 2; Length 70282;
Best Local Similarity 55.1%; Pred. No. 0.071;
Matches 98; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Oy 751 ctttatacctaactattttagtattgaagtgtagaaatacaaatltagatagta 810
    || || || || || || || || || || || || || || || || || || ||
Db 22954 CTTATATATATATATATATATATATATATATATATATATATATATATAT 23013

Oy 811 taaaatctgtctaaactctaaacraatgactagtgctaaattagccaagaagt 870
    || || || || || || || || || || || || || || || || || || ||
Db 23014 TATATATACTTTATATATATATATATATATATATATATATATATATATAT 23073

Oy 871 ttagaatggatcaactaatgttagctatacccttaactaagtattgttagat 928
    || || || || || || || || || || || || || || || || || || ||
Db 23074 ATAAATATATATATATATATATATATATATATATATATATATATATATAT 23131

RESULT 11
AC092028      78167 bp      DNA      linear      PRI 10-JAN-2002
LOCUS      AC092028
DEFINITION      Homo sapiens BAC clone RP11-527E14 from 7, complete sequence.
ACCESSION      AC092028
VERSION      AC092028.4      GI:17861068
KEYWORDS      HTG.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 78167)
AUTHORS      Sulston, D.E. and Waterston, R.
TITLE      Toward a complete human genome sequence
MEDIINE      Genome Res. 8 (11), 1097-1108 (1998)
99063792
REFERENCE      2 (bases 1 to 78167)
AUTHORS      Trani, L., Meyer, R. and Doeber, A.
TITLE      The sequence of Homo sapiens BAC clone RP11-527E14
JOURNAL      Unpublished (2002)
REFERENCE      3 (bases 1 to 78167)
AUTHORS      Waterston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (15-JUN-2001) Genome Sequencing Center, Washington
MO 63108, USA
4 (bases 1 to 78167)
REFERENCE      4 (bases 1 to 78167)
AUTHORS      Waterston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (15-DEC-2001) Genome Sequencing Center, Washington
MO 63108, USA
5 (bases 1 to 78167)
REFERENCE      5 (bases 1 to 78167)
AUTHORS      Waterston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (03-JAN-2002) Genome Sequencing Center, Washington
MO 63108, USA
6 (bases 1 to 78167)
REFERENCE      6 (bases 1 to 78167)
AUTHORS      Waterston, R.
TITLE      Direct Submission
JOURNAL      Submitted (10-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Dec 15, 2001 this sequence version replaced gi:16596649.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WTGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0527E14

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Catanesi, J.J. and de Jong, P.J. (1996) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-730B22, 2000 bp overlap; the clone sequenced to the right is RP11-147A22, 2000 bp overlap.

Polymorphisms exist between AC092032, AC093376 and AC092028.

FEATURES

```

source
1..78167
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7"
/clone="RP11-527E14"
/clone_lib="RPCT-11"
1..84
/rpt_family="Alu"
repeat_region
88..114
/rpt_family="(T)n"
repeat_region
399..633
/rpt_family="Alu"
repeat_region
635..874
/rpt_family="L1"
repeat_region
876..929
/rpt_family="Alu"
repeat_region
936..987
/rpt_family="L1"
repeat_region
988..1286
/rpt_family="Alu"
repeat_region
990..1009
/rpt_family="(TTTGG)n"
repeat_region
1154..1184
/rpt_family="AT-rich"
repeat_region
1287..1733
/rpt_family="L1"
repeat_region
2425..2703
/rpt_family="Alu"
repeat_region
3151..3339
/rpt_family="MIR"
repeat_region
3370..3667

```

```
repeat_region      /rpt_family="Alu"
3530. 3551
/rpt_family="AT_rich"
repeat_region      3679. 3702
/rpt_family="AT_rich"
repeat_region      4105. 4394
/rpt_family="Alu"
repeat_region      4396. 4455
/rpt_family="MALR"
repeat_region      4472. 4774
/rpt_family="Alu"
repeat_region      5134. 5243
/rpt_family="MIR"
repeat_region      5249. 5406
/rpt_family="Alu"
repeat_region      5483. 5794
/rpt_family="Alu"
repeat_region      5765. 5799
/rpt_family="(TAA)n"
repeat_region      6314. 6609
/rpt_family="Alu"
repeat_region      6679. 6989
/rpt_family="Alu"
repeat_region      6807. 6827
/rpt_family="AT_rich"
repeat_region      6990. 7291
/rpt_family="Alu"
repeat_region      7269. 7291
/rpt_family="(TAA)n"
repeat_region      7774. 8075
/rpt_family="Alu"
repeat_region      8090. 8389
/rpt_family="Alu"
repeat_region      8533. 8605
/rpt_family="Alu"
repeat_region      8611. 9090
/rpt_family="ERVk"
repeat_region      9176. 9458
/rpt_family="Alu"
repeat_region      9525. 9792
/rpt_family="Alu"
repeat_region      9830. 10139
/rpt_family="Alu"
repeat_region      10112. 10139
/rpt_family="Alu"
repeat_region      10507. 10691
/rpt_family="(A)n"
repeat_region      10692. 10720
/rpt_family="(T)n"
repeat_region      10996. 11384
/rpt_family="L1"
misc_feature       11365. 11587
/note="match to EST BP948153 (NID:912365428)"
repeat_region      11769. 12054
/rpt_family="Alu"
repeat_region      12154. 12293
/rpt_family="L2"
repeat_region      12404. 12671
/rpt_family="Alu"
repeat_region      12865. 13130
/rpt_family="Alu"
repeat_region      13383. 13582
/rpt_family="Alu"
repeat_region      13584. 13631
/rpt_family="Alu"
repeat_region      13634. 13726
/rpt_family="AT_rich"
repeat_region      13755. 13797
/rpt_family="AT_rich"
repeat_region      13851. 14152
/rpt_family="Alu"
repeat_region      13970. 13991
/rpt_family="AT_rich"
```

```
repeat_region      14160. 14291
/rpt_family="Alu"
repeat_region      14292. 14595
/rpt_family="Alu"
repeat_region      14596. 14749

Query Match      2.8% ; Score 49.8 ; DB 9 ; Length 78167 ;
Best Local Similarity 50.6% ; Pred. No. 0.082 ;
Matches 120 ; Conservative 0 ; Mismatches 117 ; Indels 0 ; Gaps 0 ;

QY 749 ccccttatccctaagtatgatctaatgtagtaagtgtagaataaataaatttagtag 808
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13618 CCCATCTCTACAAAGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTAAA 13677
QY 809 tataaaatcttgcctaaactcctaaactaactagctagctgtaaatagccaag 868
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13678 ATTAAAAATTTTAAAGTTTAAATTTAAATTTAAATTTTAAAGTTTAAAGTTCTTAAT 13737
QY 869 gtttagaatgatacctaattgtagctataacccttaataaagtatttaattgtagat 928
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13738 TTTAAAAATGCTTCAGCTTTTAAATTTAAATTTAAATTTTAAATTTAAATTTAAATTTAA 13797
QY 929 ggcaccaacctagctaaacccaagtaacagctagctagctagctgtagtagacagtcct 985
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13798 GTCCTTAGCACTTTAAAAATTTAAAGAGCTTTTAAATTTTAAATTTTAAACATGGGCT 13854

RESULT 12
AX251272/c 5026 bp DNA linear PAT 05-OCT-2001
LOCUS AX251272
DEFINITION Sequence 240 from Patent WO0168912.
ACCESSION AX251272
VERSION AX251272.1 GI:15984695
KEYWORDS
SOURCE
synthetic construct.
ORGANISM
artificial sequence.
REFERENCE
1 (bases 1 to 5026)
AUTHORS
Olek,A., Plødenrock,C. and Berlin,K.
TITLE
Diagnosis of diseases associated with tumor suppressor genes and
oncogenes
JOURNAL
Patent: WO 0168912-A 240 20-SEP-2001;
Epidemiology AG (DE)
FEATURES
Location/Qualifiers
source 1..5026
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1467 a 121 c 1122 g 2316 t
ORIGIN
```



```
* as soon as it is available and the accession number will
* be preserved.
1 15446: contig of 15446 bp in length
15447 15546: gap of unknown length
15547 33302: contig of 17756 bp in length
33303 33402: gap of unknown length
34403 46251: contig of 12849 bp in length
46252 46352: gap of unknown length
46352 58047: contig of 11696 bp in length
58048 58147: gap of unknown length
58148 69556: contig of 11409 bp in length
69557 69656: gap of unknown length
69657 81742: contig of 12086 bp in length
81743 81842: gap of unknown length
81843 91818: contig of 9976 bp in length
91819 91918: gap of unknown length
91919 99806: contig of 7888 bp in length
99807 111036: gap of unknown length
111037 111136: contig of 11130 bp in length
120699: contig of 9563 bp in length
120700 120799: gap of unknown length
120800 126961: contig of 6162 bp in length
126962 127061: gap of unknown length
127062 133424: contig of 6363 bp in length
133425 140427: gap of unknown length
140428 140527: gap of unknown length
140528 147998: contig of 7471 bp in length
147999 148098: gap of unknown length
148099 153891: contig of 5793 bp in length
153892 153991: gap of unknown length
153992 158879: contig of 4888 bp in length
158880 158979: gap of unknown length
158980 163973: contig of 4994 bp in length
163974 164073: gap of unknown length
164073 169658: contig of 5585 bp in length
169659 169758: gap of unknown length
169759 174422: contig of 4664 bp in length
174423 174523: gap of unknown length
174523 178125: contig of 3603 bp in length
178126 178225: gap of unknown length
178226 180622: contig of 2397 bp in length
180623 180722: gap of unknown length
180723 182915: contig of 2193 bp in length
182916 183015: gap of unknown length
183016 185168: contig of 2153 bp in length
185169 185268: gap of unknown length
185269 187384: contig of 2116 bp in length
187385 187484: gap of unknown length
187485 189790: contig of 2306 bp in length.
location/Qualifiers
1. .189790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3q"
/clone="Rp11-771N9"
BASE COUNT 59220 a 33899 c 32935 g 61177 t 2559 others
ORIGIN

Query Match 2.7%; Score 47.8; DB 2; Length 189790;
Best Local Similarity 48.4%; Pred. No. 0.31;
Matches 106; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 762 aatgataatagatcagaagcgtgagaaatcacaatttagatagataaatactctg 821
|||||
Db 126751 AATTAATATTAATTTTAAATATTAATTAATTAATTAATTAATTAATTA 126692
|||||

QY 822 ttctaactctaaactaactagctagctgctaaatagccaaaaggtttagaattg 881
|||||
Db 126691 TANNAAATATATNATTTTNNATTTTATTAATAAATATTAATTAATTA 126632
|||||
```

```
QY 882 caactaatgttagctatacccttaataagataatgttagatgagcccaactgac 941
|||||
Db 126631 AAAATATATTTTNNATTAATAAATAAATTAATTTTAATTAATTAATTA 126572
|||||

QY 942 taaacaagttacaagttagcttagatgtttagaacat 980
|||||
Db 126571 NAAATTAATAAATAATTTNNANATATNTATTTATTAATAT 126533
|||||

RESULT 15
AX346698/c AX346698 5416 bp DNA linear PAT 01-FEB-2002
LOCUS Sequence 1769 from Patent WO0200928.
DEFINITION AX346698
ACCESSION AX346698
VERSION AX346698.1 GI:18494584
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (sites)
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with the immune system
JOURNAL Patent: WO 0200928-A 1769 03-JAN-2002;
Biogenomics AG (DE)
FEATURES
Location/Qualifiers
1. .5416
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1476 a 96 c 1134 g 2710 t
ORIGIN

Query Match 2.7%; Score 47.2; DB 6; Length 5416;
Best Local Similarity 50.3%; Pred. No. 0.26;
Matches 144; Conservative 0; Mismatches 138; Indels 4; Gaps 1;

QY 715 aaagtagcaatgagcccccctcaattttagacccttatacccaatgataattag 774
|||||
Db 3899 AAATACCCCCCAATCCTACATATATCTTTTATTCCTGTAATCCGTAATCCAAA 3840
|||||

QY 775 tattaagtgtagaaaaaactcaaatltagatagataaactgtgtcnaactctca 834
|||||
Db 3839 TATTAATTTTACCAAAAGCTATTAATATCCCAATATTAATATTAATTAATAT 3780
|||||

QY 835 aaactaagctagctgctaaactagccaaaaggttagaactgataactaattgta 894
|||||
Db 3779 AATA----ATCTATTCCTTTAATCATTAATAATTAATTAATCAACAATACTT 3724
|||||

QY 895 gctatacccttaataaagcttaattgttagatgagcccaactagctaaacaagtta 954
|||||
Db 3723 TTACTAAATTTATTAATTTCTAATTTTAAAACTCTTATTAATAAATACTACTT 3664
|||||

QY 955 cagttagcttagatgtttagaacaatgagcccaacacataatcat 1000
|||||
Db 3663 TATAAAGCTACAAATTTTATTAAGCTTAATAATAATTAATTAATTA 3618
|||||
```

Search completed: July 30, 2002, 18:17:32
Job time: 5258 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2002, 16:54:10 ; Search time 219.43 Seconds
(without alignments)
13770.995 Million cell updates/sec

Title: US-09-786-835-1
Perfect score: 1760
Sequence: 1 ctcgagatgttgatagaag.....ggtcaccatcatcactcca 1760

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N_Geneseq_032802: *
2: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
3: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
4: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
5: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
6: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
7: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
8: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
9: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
10: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
11: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
12: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
13: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
14: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
15: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
16: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
17: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
18: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
19: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
20: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
21: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
22: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
23: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
24: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
25: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1760	100.0	1760	21	AAZ52234
2	1760	100.0	3416	21	AAZ52247
3	1710	97.2	1710	21	AAZ52244
4	1550	88.1	1550	21	AAZ52243
5	1497	85.1	1497	21	AAZ52242
6	1354	76.9	1354	21	AAZ52241
7	1348	76.6	1348	21	AAZ52235
8	1096	62.3	1096	21	AAZ52236
9	746	42.4	746	21	AAZ52237

10	554	31.5	554	21	AAZ52238	Maize glycine-rich
11	406	23.1	406	21	AAZ52239	Maize glycine-rich
12	190	10.8	190	21	AAZ52240	Maize glycine-rich
13	48.8	2.8	5026	22	AAZ52248	Tumour suppressor
14	48.4	2.7	7503	24	ABL33795	Human immune syste
15	47.2	2.7	5416	24	ABL33796	Human immune syste
16	45.4	2.6	6412	24	AAZ52245	Human immune syste
17	45.2	2.6	14924	22	AAZ52224	Human immune syste
18	44.8	2.5	127197	22	AAZ52246	Human immune syste
19	44.4	2.5	5452	22	AAZ52247	Human immune syste
20	44.2	2.5	17421	22	AAZ52249	Human immune syste
21	43.8	2.5	18154	24	ABL32254	Chemically pretrea
22	43.8	2.5	18997	24	ABL32570	Human immune syste
23	43.4	2.5	9510	22	AAZ52243	Tumour suppressor
24	43.4	2.5	9510	22	AAZ52243	Tumour suppressor
25	43.2	2.5	6261	22	AAZ52243	Human metastasis a
26	43.2	2.5	6261	24	AAZ52243	Human gene regulat
27	43.2	2.5	15853	24	AAZ52243	Human gene regulat
28	42.6	2.4	4590	22	AAZ52243	Human gene regulat
29	42.6	2.4	34769	22	AAZ52243	Human gene regulat
30	42.2	2.4	11726	24	ABL34063	Tumour suppressor
31	42.2	2.4	12025	24	ABL33299	Human immune syste
32	42.2	2.4	15282	24	AAZ52243	Human immune syste
33	41.6	2.4	8305	24	ABL33569	Human immune syste
34	41.6	2.4	12138	24	ABL33629	Human immune syste
35	41.6	2.4	335913	22	AAZ52243	Human immune syste
36	41.6	2.4	335913	22	AAZ52243	Human immune syste
37	41.4	2.4	335913	22	AAZ52243	Human immune syste
38	41.2	2.3	10619	22	AAZ52243	Human immune syste
39	41.2	2.3	10619	24	ABL32070	Human immune syste
40	41.2	2.3	10886	24	ABL34135	Human immune syste
41	41.2	2.3	40862	24	ABL34073	Human immune syste
42	41	2.3	6475	24	ABL34248	Human immune syste
43	41	2.3	13814	24	ABL33192	Human immune syste
44	40.8	2.3	6587	24	ABL33384	Human immune syste
45	40.4	2.3	6207	22	AAZ52243	Tumour suppressor

ALIGNMENTS

RESULT 1
AAZ52234
AAZ52234 standard; DNA; 1760 BP.
ID AAZ52234
AC AAZ52234;
XX 18-JUL-2000 (first entry)
DT
DE Maize glycine-rich protein 3 gene 5' regulatory element #1.
XX
KW Maize; glycine-rich protein 3; GRP3; 5' regulatory element;
KW root specific gene expression; root abundant gene; monooctyladen;
KW pathogen resistance; pest; herbicide; growth rate; ds.
XX
OS Zea mays.
XX
FH Key
FT 1712..1719
FT TATA_signal
FT /*tag= a
FT /standard_name= "TATA-box"
XX
XX
XX W0200015662-A1.
XX
XX 23-MAR-2000.
XX
XX 10-SEP-1999; 99WO-EP06692.
XX
XX 11-SEP-1998; 98EP-0117251.
XX
XX (AGRI-) AGRIC TECHNOLOGY & GENETICS GMBH.
XX
XX Felix G, Wulff D;

RESULT 2
AA52247
ID AA52247 standard; DNA; 3416 BP.
XX
AC AA52247;
XX
DT 18-JUL-2000 (first entry)
XX
DE Maize glycine-rich protein 3 partial genomic clone zmcGRP3.
XX
KW Maize; glycine-rich protein 3; GRP3; regulatory element; zmcGRP3;
KW root specific gene expression; root abundant gene; monocotyledon;
XX pathogen resistance; pest; herbicide; growth rate; ds.
OS
XX Zea mays. `

Key Location/Qualifiers
FT promoter 1..1780
FT /tag= a
FT TATA_signal 1712..1719
FT /tag= b
FT CDS 1781..2560
FT /tag= c
FT /product= "GRP3"
FT /note= "Glycine-rich protein 3"
FT /tag= d
FT /label= 3'-regulatory_element
FT 2692..2696
FT /tag= e
FT polyA_signal

WO20015662-A1.
XX
XX 23-MAR-2000.
XX
XX 10-SEP-1999; 99WO-EP06692.
XX
XX 11-SEP-1998; 98EP-0117251.
XX
XX (AGRI-) AGRIC TECHNOLOGY & GENETICS GMBH.
XX
XX Felix G, Wulff D;
XX
XX WPI; 2000-271382/23.
XX
XX Nucleic acid sequence is used for cloning and expressing a root
XX specific or root abundant gene in a plant -
XX
XX
XX Claim 4; Page 55-56; 60pp; English.
XX
XX The present sequence is a partial genomic clone zmcGRP3 encoding
XX maize glycine-rich protein 3 (GRP3). This sequence comprises 5' and
XX 3' regulatory elements useful for cloning and
XX expressing root specific or root abundant genes in plants, especially
XX monocots which provide high expression efficiency and high
XX tissue specificity. Root preferred gene expression provides several
XX advantages to plants e.g. resistance to pathogens, pests, herbicides and
XX adverse weather conditions, modification of growth rate and alteration of
XX root tissue function. This sequence also provides a means of isolating
XX related regulatory sequences of other plant species which confer root
XX specificity to genes of interest operably linked to them.
XX
XX Sequence 3416 BP; 755 A; 841 C; 950 G; 870 T; 0 other;

Query Match 100.0%; Score 1760; DB 21; Length 3416;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctcgagatgttgatagaagaaggaattgcgagcaaccgaggtctccaagcagcccaagc 60
|||||

Db 1 ctcgagatgttgatagaagaaggaattgcgagcaaccgaggtctccaagcagcccaagc 60
QY 61 cgcatactgcgcagtgaggtgtgtcgtgatactgtcgcgaagaagcttcgattgttcgcgat 120
|||||
Db 61 cgcatactgcgcagtgaggtgtgtcgtgatactgtcgcgaagaagcttcgattgttcgcgat 120
QY 121 cgcgagcttcagatttggtgcctccagcccaagattgttgccgcggtgtgataagc 180
|||||
Db 121 cgcgagcttcagatttggtgcctccagcccaagattgttgccgcggtgtgataagc 180
QY 181 agagccagaccccgagaccggtgtctccgataccaattgttaagaataagcagcgacaag 240
|||||
Db 181 agagccagaccccgagaccggtgtctccgataccaattgttaagaataagcagcgacaag 240
QY 241 cagcgacaacatccgagaagataagttgggggagagcagataagcttggaagaagaag 300
|||||
Db 241 cagcgacaacatccgagaagataagttgggggagagcagataagcttggaagaagaag 300
QY 301 aacaatagtgacagcgtagattctctctgtctctgtctctgttaacctctcaataat 360
|||||
Db 301 aacaatagtgacagcgtagattctctctgtctctgtctctgttaacctctcaataat 360
QY 361 actccattacaaggtcgcgcaccaacacagcgaccagtcgcgttccacgaacttggatt 420
|||||
Db 361 actccattacaaggtcgcgcaccaacacagcgaccagtcgcgttccacgaacttggatt 420
QY 421 gggttgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 480
|||||
Db 421 gggttgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 480
QY 481 gctcgcgcgtctacacatctctgtgtgttgccagacattcgggttgccagcggttaact 540
|||||
Db 481 gctcgcgcgtctacacatctctgtgtgttgccagacattcgggttgccagcggttaact 540
QY 541 gcgcctgtcctcaagaagaattgacctacacagtgctcgtcgtcgtcgtcgtcgtcgtcgtc 600
|||||
Db 541 gcgcctgtcctcaagaagaattgacctacacagtgctcgtcgtcgtcgtcgtcgtcgtcgtc 600
QY 601 gaggtcgcagtagcaggaagaagcgtgagctgtgatactgtctaccgcgttataataacg 660
|||||
Db 601 gaggtcgcagtagcaggaagaagcgtgagctgtgatactgtctaccgcgttataataacg 660
QY 661 agcttactcgtactgactaagcacaacagcgagctccagaggggggccaagta 720
|||||
Db 661 agcttactcgtactgactaagcacaacagcgagctccagaggggggccaagta 720
QY 721 ggcacatgccccccctcaatttgtacaccccttataccaatgattattagattaa 780
|||||
Db 721 ggcacatgccccccctcaatttgtacaccccttataccaatgattattagattaa 780
QY 781 ggtgaagaaatcaaatcttagatagatataaattctgtttctaactctcaaaacta 840
|||||
Db 781 ggtgaagaaatcaaatcttagatagatataaattctgtttctaactctcaaaacta 840
QY 841 atagctagttgttaaaattagccaagaagtttagaatgatacaactaatgttagctata 900
|||||
Db 841 atagctagttgttaaaattagccaagaagtttagaatgatacaactaatgttagctata 900
QY 901 ccccttaataagttatattgttagatgagcccaactgactgaacccagttaaacagta 960
|||||
Db 901 ccccttaataagttatattgttagatgagcccaactgactgaacccagttaaacagta 960
QY 961 gctctagatggttaagaacatgcccataaacaacataatttaggcctacgcgcgattct 1020
|||||
Db 961 gctctagatggttaagaacatgcccataaacaacataatttaggcctacgcgcgattct 1020
QY 1021 ctcggttcaagtcctagcatctacccagcttgactctaatctccgacaccactaaacca 1080
|||||
Db 1021 ctcggttcaagtcctagcatctacccagcttgactctaatctccgacaccactaaacca 1080
QY 1081 gctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1140
|||||
Db 1081 gctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1140

QY 1141 gccctctgtagcgcctccgtccctgtctctgttttctgtagcgtgacccgcgtgac 1200
|||||
DB 1141 gccctctgtagcgcctccgtccctgtctctgttttctgtagcgtgacccgcgtgac 1200
QY 1201 aaattgattcgtcgtcgcacagatatactatccctcggtcgtcgtgagactgacg 1260
|||||
DB 1201 aaattgattcgtcgtcgcacagatatactatccctcggtcgtcgtgagactgacg 1260
QY 1261 aactagaaattatggtctcattgtgtctttttggccctcctaatttgcctctg 1320
|||||
DB 1261 aactagaaattatggtctcattgtgtctttttggccctcctaatttgcctctg 1320
QY 1321 tcgcacactataagctgagggtgcgggtacccggtatcgtaacttgaaatgcgaacag 1380
|||||
DB 1321 tcgcacactataagctgagggtgcgggtacccggtatcgtaacttgaaatgcgaacag 1380
QY 1381 tgatgctactatactaatgtcccgtagcaataataactgttcgcgacgattggaact 1440
|||||
DB 1381 tgatgctactatactaatgtcccgtagcaataataactgttcgcgacgattggaact 1440
QY 1441 ctgattttgacagaacaaagcaactgtctatgtctataatttccagagatcga 1500
|||||
DB 1441 ctgattttgacagaacaaagcaactgtctatgtctataatttccagagatcga 1500
QY 1501 ttctcaagctgtagctgacacaatcgatgcaatttgcacagacgatatatgtatg 1560
|||||
DB 1501 ttctcaagctgtagctgacacaatcgatgcaatttgcacagacgatatatgtatg 1560
QY 1561 ttctcttaccgatacgttggttcaacagtgcgatccagatttctctgcctga 1620
|||||
DB 1561 ttctcttaccgatacgttggttcaacagtgcgatccagatttctctgcctga 1620
QY 1621 tcggatctatcgacattgacacccatcttgcacaaccaaagcgaggcggaacccgaa 1680
|||||
DB 1621 tcggatctatcgacattgacacccatcttgcacaaccaaagcgaggcggaacccgaa 1680
QY 1681 acctcgctcatgacagaccccaacgataataataccatgcatcgaatgcagcg 1740
|||||
DB 1681 acctcgctcatgacagaccccaacgataataataccatgcatcgaatgcagcg 1740
QY 1741 ggtcatcatcatcgactcca 1760
|||||
DB 1741 ggtcatcatcatcgactcca 1760

RESULT 3
AA52244
ID AA52244 standard; DNA: 1710 BP.
XX
AC AA52244;
XX
AT 18-JUL-2000 (first entry)
XX
DE Maize glycine-rich protein 3 gene 5' regulatory element #11.
XX
KW Maize; glycine-rich protein 3; GRP3; 5' regulatory element;
KM root specific gene expression; root abundant gene; monocotyledon;
KW pathogen resistance; pest; herbicide; growth rate; ds.
XX
OS Zea mays.
XX
OS
PN MO200015662-A1.
XX
PD 23-MAR-2000.
XX
PF 10-SEP-1999; 99WO-EP06692.
XX
PR 11-SEP-1998; 98EP-0117251.
XX
PA (AGRI-) AGRIC TECHNOLOGY & GENETICS GMBH.
XX
PI Felix G, Wulff D;

XX
DR WPI: 2000-271382/23.
XX
PT Nucleic acid sequence is used for cloning and expressing a root
PS specific or root abundant gene in a plant -
XX
XX
Claim 3; Page 54; 60pp; English.
CC The present sequence is a 5' regulatory element of maize glycine-rich
CC protein 3 (GRP3) gene. It corresponds to positions 1-1710 of the
CC partial genomic clone zmgRP3. The present sequence is useful for cloning
CC and expressing root specific or root abundant genes in plants,
CC especially monocots which provide high expression efficiency and high
CC tissue specificity. Root preferred gene expression provides several
CC advantages to plants e.g. resistance to pathogens, pests, herbicides and
CC adverse weather conditions, modification of growth rate and alteration of
CC root tissue function. This sequence also provides a means of isolating
CC related regulatory sequences of other plant species which confer root
CC specificity to genes of interest operably linked to them.
XX
XX
SO Sequence 1710 BP; 417 A; 446 C; 406 G; 441 T; 0 other;

Query Match 97.2%; Score 1710; DB 21; Length 1710;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1710; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctgagatgttggatagaagaaggaattgcagacacgaggtcccaagcgacgcaacg 60
|||||
DB 1 ctgagatgttggatagaagaaggaattgcagacacgaggtcccaagcgacgcaacg 60
QY 61 cgcacctcgcagtgaggtgtgtcgtatgtcgcgaagaaggtctgattgtctgcgat 120
|||||
DB 61 cgcacctcgcagtgaggtgtgtcgtatgtcgcgaagaaggtctgattgtctgcgat 120
QY 121 cgcgagcttcagatttggctcattgacacccagattgttggcggggcggtgataagc 180
|||||
DB 121 cgcgagcttcagatttggctcattgacacccagattgttggcggggcggtgataagc 180
QY 181 agacgcgaacccgagacccggtgtctccgataccaattgttagaataagcgcgacaa 240
|||||
DB 181 agacgcgaacccgagacccggtgtctccgataccaattgttagaataagcgcgacaa 240
QY 241 cagcgaacaatccgagagacatagtttgggggagagacagatcttggagagaagaag 300
|||||
DB 241 cagcgaacaatccgagagacatagtttgggggagagacagatcttggagagaagaag 300
QY 301 aacatagtgacagctagattctctctgttccctctgttaccctctccaatctaat 360
|||||
DB 301 aacatagtgacagctagattctctctgttccctctgttaccctctccaatctaat 360
QY 361 actccattacaaggtcggcccatcagcgacccagtcgcggttccacgaacttggatt 420
|||||
DB 361 actccattacaaggtcggcccatcagcgacccagtcgcggttccacgaacttggatt 420
QY 421 gggttgtgcgagcgcgacgtgtcttgcctcgtcgtggaatctgttcttggagctgt 480
|||||
DB 421 gggttgtgcgagcgcgacgtgtcttgcctcgtcgtggaatctgttcttggagctgt 480
QY 481 gctcggctgttaacaactcttcttgttgcgaagaattcgcgtgacacgcttaact 540
|||||
DB 481 gctcggctgttaacaactcttcttgttgcgaagaattcgcgtgacacgcttaact 540
QY 541 ggcctctgctacaagaattgctctataacagtcctgtcgtcagctgcagcgcttgaatcc 600
|||||
DB 541 ggcctctgctacaagaattgctctataacagtcctgtcgtcagctgcagcgcttgaatcc 600
QY 601 gagtgcgcagtagaggaacagctgagctgcatagcctgtaccggcttataataactg 660
|||||
DB 601 gagtgcgcagtagaggaacagctgagctgcatagcctgtaccggcttataataactg 660
QY 661 agcttactgctcagcttagctatagcaccagaagcggaaccagaagggggcaaatla 720
|||||

Dh 661 agcttactcgtaagtagctatagcgaccagagcgatccagaggggagcaagta 720
QY 721 ggcattggccccctcaatttggtaaccctttatatacccaatgattattagattaa 780
Dh 721 ggcattggccccctcaatttggtaaccctttatatacccaatgattattagattaa 780
QY 781 ggtagaaaaatacaaatltagatataaaaaatctgttccaaatctctaaacta 840
Dh 781 ggtagaaaaatacaaatltagatataaaaaatctgttccaaatctctaaacta 840
QY 841 atagctgtgtcctaaatattagccaaaaagtttagaattgataactaattgttagctata 900
Dh 841 atagctgtgtcctaaatattagccaaaaagtttagaattgataactaattgttagctata 900
QY 901 ccccttaataagattattattgttagagggcccaaccagtaaacagtttaaacgtta 960
Dh 901 ccccttaataagattattattgttagagggcccaaccagtaaacagtttaaacgtta 960
QY 961 gctctagatgtttagaacatggcctaaacaacacatattatagacctactgcgcatctc 1020
Dh 961 gctctagatgtttagaacatggcctaaacaacacatattatagacctactgcgcatctc 1020
QY 1021 ctggttcagtcctcagatcctaacagtcctgaacttaactcctccgaacacccaataacca 1080
Dh 1021 ctggttcagtcctcagatcctaacagtcctgaacttaactcctccgaacacccaataacca 1080
QY 1081 gtcgtcgcgtgcgtcgttctccgcacacgacacacatcccgatgtagaagtaaccgacttc 1140
Dh 1081 gtcgtcgcgtgcgtcgttctccgcacacgacacacatcccgatgtagaagtaaccgacttc 1140
QY 1141 gccctctgtagcgctccgtccctgtccctgttcttctcgtagacgtcgcgtcgcac 1200
Dh 1141 gccctctgtagcgctccgtccctgtccctgttcttctcgtagacgtcgcgtcgcac 1200
QY 1201 aaatagatctgctcgtcgcgaagaatagactataccctcgtcgtcgtcgtgtagcttaacg 1260
Dh 1201 aaatagatctgctcgtcgcgaagaatagactataccctcgtcgtcgtcgtgtagcttaacg 1260
QY 1261 aactagaataatagctcctcatgtgtcttcttggccctcccaaatlcttcgcgtcgcg 1320
Dh 1261 aactagaataatagctcctcatgtgtcttcttggccctcccaaatlcttcgcgtcgcg 1320
QY 1321 tcgcgcactctatagctgcgggagtcgagggtacgtaatcgtgaatcgtgaatcgtgaacag 1380
Dh 1321 tcgcgcactctatagctgcgggagtcgagggtacgtaatcgtgaatcgtgaatcgtgaacag 1380
QY 1381 tgaatgagtaactataatcccgtagcaataataatacagttgtagcagatgggaatct 1440
Dh 1381 tgaatgagtaactataatcccgtagcaataataatacagttgtagcagatgggaatct 1440
QY 1441 ctgattttgacagaagaacgaacgaactgctatgctagctaatcaattccagagagatcga 1500
Dh 1441 ctgattttgacagaagaacgaacgaactgctatgctagctaatcaattccagagagatcga 1500
QY 1501 ttctctacagtgctcagtgcaacaaatcgatgcaattgtagcagacgatatagtgtaatgg 1560
Dh 1501 ttctctacagtgctcagtgcaacaaatcgatgcaattgtagcagacgatatagtgtaatgg 1560
QY 1561 ttctcttattcagatcgtgtgtaacagtgtagatccagattacagagatttctctgcgcttga 1620
Dh 1561 ttctcttattcagatcgtgtgtaacagtgtagatccagattacagagatttctctgcgcttga 1620
QY 1621 tcggatctcatcgcacatggagacccatctgccaacccaacgcggggcgaggaacccgaa 1680
Dh 1621 tcggatctcatcgcacatggagacccatctgccaacccaacgcggggcgaggaacccgaa 1680
QY 1681 acatgcgctcagatgcagaccccaacgag 1710
Dh 1681 acatgcgctcagatgcagaccccaacgag 1710

RESULT 4
AA52243

ID AA52243 standard: DNA; 1550 BP.
XX
AC AA52243;
XX
DT 18-JUL-2000 (first entry)
XX
DE Maize glycine-rich protein 3 gene 5' regulatory element #10.
XX
KW Maize; glycine-rich protein 3; GRP3; 5' regulatory element;
RV root specific gene expression; root abundant gene; monocotyledon;
XX pathogen resistance; pest; herbicide; growth rate; ds.
OS Zea mays.
XX
XX W0200015662-A1.
XX
XX 23-MAR-2000.
XX
XX 10-SEP-1999; 99WO-EP06692.
XX
XX 11-SEP-1998; 98EP-0117251.
XX
XX (AGRI-) AGRIC TECHNOLOGY & GENETICS GMBH.
XX
XX Felix G. Wulff D;
XX
XX WPI; 2000-271382/23.
XX
XX Nucleic acid sequence is used for cloning and expressing a root
PT specific or root abundant gene in a plant -
XX
XX Claim 3; Page 53-54; 60pp; English.
XX
XX The present sequence is a 5' regulatory element of maize glycine-rich
CC protein 3 (GRP3) gene. It corresponds to positions 1-1550 of the
CC partial genomic clone zmGRP3. The present sequence is useful for cloning
CC and expressing root specific or root abundant genes in plants,
CC especially monocots which provide high expression efficiency and high
CC tissue specificity. Root preferred gene expression provides several
CC advantages to plants e.g. resistance to pathogens, pests, herbicides and
CC adverse weather conditions, modification of growth rate and alteration of
CC root tissue function. This sequence also provides a means of isolating
CC related regulatory sequences of other plant species which confer root
CC specificity to genes of interest operably linked to them.
XX
SQ Sequence 1550 BP; 379 A; 394 C; 369 G; 408 T; 0 other;

Query Match 88.1%; Score 1550; DB 21; Length 1550;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1550; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctgagatgttgaatagaagaagaattgcgagcaacgaggtctccaagcgacaacg 60
Dh 1 ctgagatgttgaatagaagaagaattgcgagcaacgaggtctccaagcgacaacg 60
QY 61 cgaactcgtgcagtgagtggtgctgctgtagtcgcgaagaagagcttgatgttcgat 120
Dh 61 cgaactcgtgcagtgagtggtgctgctgtagtcgcgaagaagagcttgatgttcgat 120
QY 121 cgcagatcagatttggctccatgacccagagattgttggcgcgcggtgataagc 180
Dh 121 cgcagatcagatttggctccatgacccagagattgttggcgcgcggtgataagc 180
QY 181 agacgcagaccgcgagacgggtgtctccgatalaccaatgtttagcaatagcagcgacaag 240
Dh 181 agacgcagaccgcgagacgggtgtctccgatalaccaatgtttagcaatagcagcgacaag 240
QY 241 cagcgacaatccgagagatagtttggggggagcaatagcttgggaggaagaag 300
Dh 241 cagcgacaatccgagagatagtttggggggagcaatagcttgggaggaagaag 300
QY 301 aacatagtgacagcgtagattctctctgtctcctctgttaacctctcaatcaat 360

```

Db      301 aacaaatagtagacgtagatccctctcgtgctcctcgttaacctctcctcaatcaat 360
Qy      361 actccatatacaaggctcgcccatcagcgcaaccagtcgaggtccacgaacttggagtt 420
Db      361 actccatatacaaggctcgcccatcagcgcaaccagtcgaggtccacgaacttggagtt 420
Qy      421 ggggttgtagcagcgcgcgacttggtcttgctgtgtgaggtctgtctgtgaggtct 480
Db      421 ggggttgtagcagcgcgcgacttggtcttgctgtgtgaggtctgtctgtgaggtct 480
Qy      481 gctcgcgcgcaacaacatctctgtgttgtagcagcactcgcgtgcaacgcgcttaact 540
Db      481 gctcgcgcgcaacaacatctctgtgttgtagcagcactcgcgtgcaacgcgcttaact 540
Qy      541 ggcgtctgtccaagaanaatgtccatatacagtcgcgtgctcagctcagcgcttgaatcc 600
Db      541 ggcgtctgtccaagaanaatgtccatatacagtcgcgtgctcagctcagcgcttgaatcc 600
Qy      601 gaggtcgcaagtagcagcgcaaacgctgagctgcatgcttaaccggtcttaataactg 660
Db      601 gaggtcgcaagtagcagcgcaaacgctgagctgcatgcttaaccggtcttaataactg 660
Qy      661 agctctaccgcgtacgcttagctatagcgaccagagcgagatccagaggggggcaagta 720
Db      661 agctctaccgcgtacgcttagctatagcgaccagagcgagatccagaggggggcaagta 720
Qy      721 ggcacatgccccccctcaattttgtacacccctttatacctaattgattattatgattaa 780
Db      721 ggcacatgccccccctcaattttgtacacccctttatacctaattgattattatgattaa 780
Qy      781 ggtgtagaaaaatacaaatatttagatagtagtaaaaatctgtctcaaatctcaaaaacta 840
Db      781 ggtgtagaaaaatacaaatatttagatagtagtaaaaatctgtctcaaatctcaaaaacta 840
Qy      841 atagctagttgtcaaaatttagccaanaagtttagaagtgcataactaatgttagctata 900
Db      841 atagctagttgtcaaaatttagccaanaagtttagaagtgcataactaatgttagctata 900
Qy      901 ccccttaaaaataatattatgtttagatgcccacacttaagctaaacgaattacaattta 960
Db      901 ccccttaaaaataatattatgtttagatgcccacacttaagctaaacgaattacaattta 960
Qy      961 gctctagatgttttagaacaatgacctaaacaacatattatgagcctactcgcgcgattcct 1020
Db      961 gctctagatgttttagaacaatgacctaaacaacatattatgagcctactcgcgcgattcct 1020
Qy      1021 ctcggttcagtccttagacatctaccagtcgactctaattctcccgacaacactaaacca 1080
Db      1021 ctcggttcagtccttagacatctaccagtcgactctaattctcccgacaacactaaacca 1080
Qy      1081 gttctgcgtgctgcctgtttccgcgcaacgacccatcccgatgagatgccgacttcc 1140
Db      1081 gttctgcgtgctgcctgtttccgcgcaacgacccatcccgatgagatgccgacttcc 1140
Qy      1141 gccctctgtagcgctccgctccctgtctctgttttctcgtgtagcgtcgatccgctgac 1200
Db      1141 gccctctgtagcgctccgctccctgtctctgttttctcgtgtagcgtcgatccgctgac 1200
Qy      1201 aaattagatctcgtctgtagcaagaatagatatacctcctgcggtgaggttgaagc 1260
Db      1201 aaattagatctcgtctgtagcaagaatagatatacctcctgcggtgaggttgaagc 1260
Qy      1261 aactagaanaattatggtctcactgtgtcttttggccctcccaaaatttctgcctgcg 1320
Db      1261 aactagaanaattatggtctcactgtgtcttttggccctcccaaaatttctgcctgcg 1320
Qy      1321 tccgcacatctatagctgcgggagtgcggggtaccgtaacgtgaactctgaagtccaacag 1380
Db      1321 tccgcacatctatagctgcgggagtgcggggtaccgtaacgtgaactctgaagtccaacag 1380
Qy      1381 tgatgggtactatctatgtcccggtgcagtaataactgttgcgagcgttgggaatct 1440

```

```

Db      1381 tgatgggtactatctaatgtcccgtagcaataatcaactgtgtccgacgttggaaatct 1440
Qy      1441 ctaggttttaacgaanaacgaaggaactgtctatgtagtaataatccaagagatcga 1500
Db      1441 ctaggttttaacgaanaacgaaggaactgtctatgtagtaataatccaagagatcga 1500
Qy      1501 ttctcaagtgctagctgcacaacatcgaatgcgaatttggcatcagaagat 1550
Db      1501 ttctcaagtgctagctgcacaacatcgaatgcgaatttggcatcagaagat 1550

```

```

RESULT 5
AA52242
ID AA52242 standard; DNA; 1497 BP.
XX
AC AA52242;
XX
DT 18-JUL-2000 (first entry)
XX
DE Maize glycine-rich protein 3 gene 5' regulatory element #9.
XX
KW Maize; glycine-rich protein 3; GRP3; 5' regulatory element;
KW root specific gene expression; root abundant gene; monocotyledon;
KW pathogen resistance; pest; herbicide; growth rate; ds.
XX
OS Zea mays.
XX
PN MO200015662-A1.
XX
PD 23-MAR-2000.
XX
PF 10-SEP-1999; 99MO-EP06692.
XX
PR 11-SEP-1998; 98EP-0117251.
XX
PA (AGRI-) AGRIC TECHNOLOGY & GENETICS GMBH.
XX
PI Feix G, Muller D;
XX
DR WPI; 2000-271382/23.
XX
PT Nucleic acid sequence is used for cloning and expressing a root
PT specific or root abundant gene in a plant -
XX
PS Claim 3; Page 53; 60pp; English.
XX
CC The present sequence is a 5' regulatory element of maize glycine-rich
CC protein 3 (GRP3) gene. It corresponds to positions 1-1497 of the
CC partial genomic clone zmGRP3. The present sequence is useful for cloning
CC and expressing root specific or root abundant genes in plants,
CC especially monocots which provide high expression efficiency and high
CC tissue specificity. Root preferred gene expression provides several
CC advantages to plants e.g. resistance to pathogens, pests, herbicides and
CC adverse weather conditions, modification of growth rate and alteration of
CC root tissue function. This sequence also provides a means of isolating
CC related regulatory sequences of other plant species which confer root
CC specificity to genes of interest operably linked to them.
XX
SQ Sequence 1497 BP; 363 A; 382 C; 358 G; 394 T; 0 other:

```

```

Query Match      85.1%; Score 1497; DB 21; Length 1497;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 ctgagatgttgatagagaagaattcgagcaacgaggtctccaagcgacgccaacg 60
Db      1 ctgagatgttgatagagaagaattcgagcaacgaggtctccaagcgacgccaacg 60
Qy      61 cgcatacctcgcaatgaggtgtgtcgatgctcggaagagcttgattgtgtgat 120
Db      61 cgcatacctcgcaatgaggtgtgtcgatgctcggaagagcttgattgtgtgat 120

```

QY 121 cgcgagcttcagatttgctccatgaccccaagatctgtttgscgscgctgtlaatgagc 180
DB 121 cgcgagcttcagatttgctccatgaccccaagatctgtttgscgscgctgtlaatgagc 180
QY 181 agagcccgagcccgagaccggtgtctccgataccaattggttagcaatagcaagcaag 240
DB 181 agagcccgagcccgagaccggtgtctccgataccaattggttagcaatagcaagcaag 240
QY 241 cagcgaaacatccgagacacatagtttgggagagagcagatagtttggagggagagag 300
DB 241 cagcgaaacatccgagacacatagtttgggagagagcagatagtttggagggagagag 300
QY 301 aacaatagtgacagctgatttcctctctgtctctgttaccctctctcaatctaat 360
DB 301 aacaatagtgacagctgatttcctctctgtctctgttaccctctctcaatctaat 360
QY 361 actcattacaaagctgcgcccataagcgacccacgacccgctgcggttccacgaacttggatt 420
DB 361 actcattacaaagctgcgcccataagcgacccacgacccgctgcggttccacgaacttggatt 420
QY 421 gggatttgscgagccgacactgcttctgtctgtgagatctgtctgttggagcttt 480
DB 421 gggatttgscgagccgacactgcttctgtctgtgagatctgtctgttggagcttt 480
QY 481 gctcgcgctgtcaacaatctctgttgttgcagacactcgcgtgscagcgcttaact 540
DB 481 gctcgcgctgtcaacaatctctgttgttgcagacactcgcgtgscagcgcttaact 540
QY 541 ggcgtctgcataaagaattgctctatacagtcgctgscgttagctgcagcgcttgaatcc 600
DB 541 ggcgtctgcataaagaattgctctatacagtcgctgscgttagctgcagcgcttgaatcc 600
QY 601 gaagtcgcagtagcagagcaaacgctgacgtgcatactgtctataccgcttataataactg 660
DB 601 gaagtcgcagtagcagagcaaacgctgacgtgcatactgtctataccgcttataataactg 660
QY 661 agctctactgcgttagctatagtagcagcagagcgcgatccagaggggggcaaatga 720
DB 661 agctctactgcgttagctatagtagcagcagagcgcgatccagaggggggcaaatga 720
QY 721 ggcacatggccccctcaattttgtacacccctttatacccaatgatttagatttaa 780
DB 721 ggcacatggccccctcaattttgtacacccctttatacccaatgatttagatttaa 780
QY 781 gtgtagaaaaatacaaaattttagatagtagtaaaaatctgttctaaatctcaaaacta 840
DB 781 gtgtagaaaaatacaaaattttagatagtagtaaaaatctgttctaaatctcaaaacta 840
QY 841 atagctagttgctaaatttagccaaaaggttagaagtagcaactaatgtttagctata 900
DB 841 atagctagttgctaaatttagccaaaaggttagaagtagcaactaatgtttagctata 900
QY 901 cccttaataatgatttaattgttagatgcccacactagctaaacacgaagtaacagta 960
DB 901 cccttaataatgatttaattgttagatgcccacactagctaaacacgaagtaacagta 960
QY 961 gctctagaattttagaacaatgacctaaacaacatataatagcctactcgcggaatcct 1020
DB 961 gctctagaattttagaacaatgacctaaacaacatataatagcctactcgcggaatcct 1020
QY 1021 ctgcgttcacatccagatcctacacagctgacttaattctccgacacacccaaacca 1080
DB 1021 ctgcgttcacatccagatcctacacagctgacttaattctccgacacacccaaacca 1080
QY 1081 gctcgtcgtgcctgttttcgcgacacgacccatcccgatgagcagctcaccgacttc 1140
DB 1081 gctcgtcgtgcctgttttcgcgacacgacccatcccgatgagcagctcaccgacttc 1140
QY 1141 ggcocctcgtagcgccctcgtctcctgtctgttttctcgtgacgtcgatccgctgac 1200
DB 1141 ggcocctcgtagcgccctcgtctcctgtctgttttctcgtgacgtcgatccgctgac 1200
QY 1201 aaatlagatctcgtctgcgaacgatatgactatccctcgtcgtcggtgagattgagc 1260

DB 1201 aaatlagatctcgtctgcgaacgatatgactatccctcgtcgtcggtgagattgagc 1260
QY 1261 aactgaaatatgctctcaltggtctcttggccctcctcaaatcttgcctcgc 1320
DB 1261 aactgaaatatgctctcaltggtctcttggccctcctcaaatcttgcctcgc 1320
QY 1321 tccgcacacttagctgtgaggagtgcggggtacccgtatcgtgaatctgaaatcgacag 1380
DB 1321 tccgcacacttagctgtgaggagtgcggggtacccgtatcgtgaatctgaaatcgacag 1380
QY 1381 tgatggcgtacatctaatgctccgtgcagtaataatcactgttgcgcagatgggaatct 1440
DB 1381 tgatggcgtacatctaatgctccgtgcagtaataatcactgttgcgcagatgggaatct 1440
QY 1441 ctgattttgacagaacaaagcgaactgctatgctagctaatataatccagagagat 1497
DB 1441 ctgattttgacagaacaaagcgaactgctatgctagctaatataatccagagagat 1497

RESULT 6

AAZ52241 standard; DNA; 1354 BP.

AAZ52241;

18-JUL-2000 (first entry)

Maize glycine-rich protein 3 gene 5' regulatory element #8.

Maize; glycine-rich protein 3; GRP3; 5' regulatory element;

root specific gene expression; root abundant gene; monocolyledon;

pathogen resistance; pest; herbicide; growth rate; ds.

Zea mays.

W0200015662-A1.

23-MAR-2000.

10-SEP-1999; 99WO-EP06692.

11-SEP-1998; 98EP-0117251.

(AGRI-) AGRIC TECHNOLOGY & GENETICS GMBH.

Felix G. Wulff D;

Nucleic acid sequence is used for cloning and expressing a root specific or root abundant gene in a plant -

Claim 3; Page 52; 60pp; English.

The present sequence is a 5' regulatory element of maize glycine-rich protein 3 (GRP3) gene. It corresponds to positions 1-1354 of the partial genomic clone zmGRP3. The present sequence is useful for cloning and expressing root specific or root abundant genes in plants, especially monocots which provide high expression efficiency and high tissue specificity. Root preferred gene expression provides several advantages to plants e.g. resistance to pathogens, pests, herbicides and adverse weather conditions, modification of growth rate and alteration of root tissue function. This sequence also provides a means of isolating related regulatory sequences of other plant species which confer root specificity to genes of interest operably linked to them.

Sequence 1354 BP; 321 A; 353 C; 325 G; 355 T; 0 other;

Query Match 76.9%; Score 1354; DB 21; Length 1354;

Best Local Similarity 100.0%; Pred. No. 0; Matches 1354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctcgaatgttctgatatgaagaagaattgctgagcaaacgagctctccaagcagcccaag 60
 |||||
 Db 1 ctcgagatgttctgatatgaagaagaattgctgagcaaacgagctctccaagcagcccaag 60
 |||||
 QY 61 cgcatactgcgcagctgagctgctgcgtctgctgcgaagaagctctcatgttctgcgat 120
 |||||
 Db 61 cgcatactgcgcagctgagctgctgcgtctgctgcgaagaagctctcatgttctgcgat 120
 |||||
 QY 121 cgcgaagcttcagatcttgcgtccatgacccagatgttctgcgcgcgcgtctgtaataagc 180
 |||||
 Db 121 cgcgaagcttcagatcttgcgtccatgacccagatgttctgcgcgcgcgtctgtaataagc 180
 |||||
 QY 181 agcgcgcgaagcccgagacgcgtctctccgaataccaattgttagcaatagacgcgcgaag 240
 |||||
 Db 181 agcgcgcgaagcccgagacgcgtctctccgaataccaattgttagcaatagacgcgcgaag 240
 |||||
 QY 241 cagcgaacaatccgagagacatagttctgaggagagacagatgcttgcgaaggaag 300
 |||||
 Db 241 cagcgaacaatccgagagacatagttctgaggagagacagatgcttgcgaaggaag 300
 |||||
 QY 301 aacaaatagtcagacgctagatctctctctgctctctcttaccctctctcaatctaat 360
 |||||
 Db 301 aacaaatagtcagacgctagatctctctctgctctctcttaccctctctcaatctaat 360
 |||||
 QY 361 accctctatcaagctgagctgagccatcaagcgacccagctgcgcgtctccacgaactgggat 420
 |||||
 Db 361 accctctatcaagctgagctgagccatcaagcgacccagctgcgcgtctccacgaactgggat 420
 |||||
 QY 421 gggcttctgagcagcgcgactgctctctgctctgctgcgtgcgtctgtctgtgcgtg 480
 |||||
 Db 421 gggcttctgagcagcgcgactgctctctgctctgctgcgtgcgtctgtctgtgcgtg 480
 |||||
 QY 481 gctcgcgcgtcttaacaatctctctgttcttgcgaagcacttcggtgcgaagcgttact 540
 |||||
 Db 481 gctcgcgcgtcttaacaatctctctgttcttgcgaagcacttcggtgcgaagcgttact 540
 |||||
 QY 541 ggcctctgcacaagaatctgctctatcaatgctcctgcgtcgtctcagctcagcgcttgaatcc 600
 |||||
 Db 541 ggcctctgcacaagaatctgctctatcaatgctcctgcgtcgtctcagctcagcgcttgaatcc 600
 |||||
 QY 601 gaggtcgcagctagcagagcaaacgctgagctgcatgctgtacgcggtcttataactg 660
 |||||
 Db 601 gaggtcgcagctagcagagcaaacgctgagctgcatgctgtacgcggtcttataactg 660
 |||||
 QY 661 agcctactgcgtacgcttagctatagcacaagcgcgatccagagaggggagcaagta 720
 |||||
 Db 661 agcctactgcgtacgcttagctatagcacaagcgcgatccagagaggggagcaagta 720
 |||||
 QY 721 ggcataggccccctcaatttctacaccccttataccaatgatatattaglataa 780
 |||||
 Db 721 ggcataggccccctcaatttctacaccccttataccaatgatatattaglataa 780
 |||||
 QY 781 gctgtagaaaaatacaaaatttagatagatataaaatctgttcttaactccttaacta 840
 |||||
 Db 781 gctgtagaaaaatacaaaatttagatagatataaaatctgttcttaactccttaacta 840
 |||||
 QY 841 ataagctgttgcataaattagcacaaggtttagaattgatactcaactaattttagacta 900
 |||||
 Db 841 ataagctgttgcataaattagcacaaggtttagaattgatactcaactaattttagacta 900
 |||||
 QY 901 ccccttaaatagatattaattgttagatgcccacaactagctaaanaacagttaacgta 960
 |||||
 Db 901 ccccttaaatagatattaattgttagatgcccacaactagctaaanaacagttaacgta 960
 |||||
 QY 961 gctctagatgtttagaacaatgctcctaacaacacatataatgagctactgcgcgatcct 1020
 |||||
 Db 961 gctctagatgtttagaacaatgctcctaacaacacatataatgagctactgcgcgatcct 1020
 |||||
 QY 1021 ctggttcaagctcagacatctacacgctgactcaatctcccgacacacacttaacaa 1080
 |||||
 Db 1021 ctggttcaagctcagacatctacacgctgactcaatctcccgacacacacttaacaa 1080
 |||||

QY 1081 gctcgtcgtctgcctctgttcttcgcgcaacgacccatctccgatgtagcgtacccgacttc 1140
 |||||
 Db 1081 gctcgtcgtctgcctctgttcttcgcgcaacgacccatctccgatgtagcgtacccgacttc 1140
 |||||
 QY 1141 gccctctcgtagcgcctcgcctctctctgttcttctctgtgtagcgtcgcgtcgcgac 1200
 |||||
 Db 1141 gccctctcgtagcgcctcgcctctctctgttcttctctgtgtagcgtcgcgtcgcgac 1200
 |||||
 QY 1201 aaatagatctgcgtctgcgaacgatatactatccctgcgtcgtctgtagttagc 1260
 |||||
 Db 1201 aaatagatctgcgtctgcgaacgatatactatccctgcgtcgtctgtagttagc 1260
 |||||
 QY 1261 aactagaanaattatgctctcatgttgcttcttcttgcgcctcctcaaatcttctcgtcg 1320
 |||||
 Db 1261 aactagaanaattatgctctcatgttgcttcttcttgcgcctcctcaaatcttctcgtcg 1320
 |||||
 QY 1321 tcgcacactctatagctgcgcggagtgcggggtac 1354
 |||||
 Db 1321 tcgcacactctatagctgcgcggagtgcggggtac 1354
 |||||

RESULT 7

AAZ52235 standard; DNA: 1348 BP.

AAZ52235;

18-JUL-2000 (first entry)

Maize glycine-rich protein 3 gene 5' regulatory element #2.

Maize; glycine-rich protein 3; GRP3; 5' regulatory element;

KW root specific gene expression; root abundant gene; monocotyledon;

KW pathogen resistance; pest; herbicide; growth rate; ds.

OS Zea mays.

FH Key Location/Qualifiers

FT TATA_signal 1300..1307

FT /*tag= a /standard_name= "TATA-box"

FN WO200015662-A1.

PD 23-MAR-2000.

XX 10-SEP-1999; 99WO-EP06692.

XX 11-SEP-1998; 98EP-0117251.

XX (AGRI-) AGRIC TECHNOLOGY & GENETICS GMBH.

XX Feix G, Wulff D;

XX WPI: 2000-271382/23.

XX Claim 2; Page 50; 60pp; English.

CC The present sequence is a 5' regulatory element of maize glycine-rich
 CC protein 3 (GRP3) gene. It corresponds to positions 413-1760 of the
 CC partial genomic clone zmgRP3. The present sequence is useful for cloning
 CC and expressing root specific or root abundant genes in plants,
 CC especially monocots which provide high expression efficiency and high
 CC tissue specificity. Root preferred gene expression provides several
 CC advantages to plants e.g. resistance to pathogens, pests, herbicides and
 CC adverse weather conditions, modification of growth rate and alteration of
 CC root tissue function. This sequence also provides a means of isolating
 CC related regulatory sequences of other plant species which confer root
 CC specificity to genes of interest operably linked to them.

SQ Sequence 1348 BP; 331 A; 355 C; 296 G; 366 T; 0 other;

Query Match 76.6%; Score 1348; DB 21; Length 1348;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 413 ttggattggtttgttgcgacgcgactggtcttgctcgtcgtggaattgtctgtg 472
DB 1 ttggattggtttgttgcgacgcgactggtcttgctcgtcgtggaattgtctgtg 60
OY 473 gagctgtgctcgcgcgtctaaacatctctgttgttgcgacgactcgcgtggccacg 532
DB 61 gagctgtgctcgcgcgtctaaacatctctgttgttgcgacgactcgcgtggccacg 120
OY 533 cgttacgcgcgtcgtcgtcgaagaattgctcctacagtgcttgctgctgaagcg 592
DB 121 cgttacgcgcgtcgtcgtcgaagaattgctcctacagtgcttgctgctgaagcg 180
OY 593 ttgaatccgaggtcgcagtaagcaagcgtgacgtgacgtgctgtacccgcttatt 652
DB 181 ttgaatccgaggtcgcagtaagcaagcgtgacgtgacgtgctgtacccgcttatt 240
OY 653 aataactagctctactcgtcgtacgttagctaaagcaagcgcgatccagagggg 712
DB 241 aataactagctctactcgtcgtacgttagctaaagcaagcgcgatccagagggg 300
OY 713 gcaaggtgagccatgccccctcaattttgtacacccctttatacctaattatt 772
DB 301 gcaaggtgagccatgccccctcaattttgtacacccctttatacctaattatt 360
OY 773 agtataagtgtagaaaaaatcaaatatttagatagatlaaaaaatctgttctaactc 832
DB 361 agtataagtgtagaaaaaatcaaatatttagatagatlaaaaaatctgttctaactc 420
OY 833 taaaactaatagctagtgtgctaaatattagcacaaggtttagaatggtacaaactaattgt 892
DB 421 taaaactaatagctagtgtgctaaatattagcacaaggtttagaatggtacaaactaattgt 480
OY 893 tagctataccctttaaatgaatgaatgaatgttagatggtgcccccaactagtaaaacaggt 952
DB 481 tagctataccctttaaatgaatgaatgaatgttagatggtgcccccaactagtaaaacaggt 540
OY 953 aacaggtagctctagatgttttagaacatggtcctaacaacacatattatgactactcgc 1012
DB 541 aacaggtagctctagatgttttagaacatggtcctaacaacacatattatgactactcgc 600
OY 1013 cggatccctcgtgctcagctctcagcatctacagctcgtacttaacttcccgacacac 1072
DB 601 cggatccctcgtgctcagctctcagcatctacagctcgtacttaacttcccgacacac 660
OY 1073 ctaaacccagctcgtcgtcgtccctgttcccgacccgacccaattccgattgagctcag 1132
DB 661 ctaaacccagctcgtcgtcgtccctgttcccgacccgacccaattccgattgagctcag 720
OY 1133 cggacttcgccccctcgttagcgcctcgcgtccctgttctgttcttcgtgacgtcagtc 1192
DB 721 cggacttcgccccctcgttagcgcctcgcgtccctgttctgttcttcgtgacgtcagtc 780
OY 1193 cgttcgacaattatgattcgtcgttcgcaagatagatactaccctcgtcgtcgtg 1252
DB 781 cgttcgacaattatgattcgtcgttcgcaagatagatactaccctcgtcgtcgtg 840
OY 1253 acttcgacaactagaataattatgctcctcaatgtgtcttcttggccccctcaaatcttc 1312
DB 841 acttcgacaactagaataattatgctcctcaatgtgtcttcttggccccctcaaatcttc 900
OY 1313 gtccctggtccgcaactctatagctcgggaggtgcggggtacggtatcgtaactgga 1372
DB 901 gtccctggtccgcaactctatagctcgggaggtgcggggtacggtatcgtaactgga 960
OY 1373 tcgaacagtgtgtggttctatctatgtcccggtcagtaataacagttgcgacgatt 1432

DB 961 tcgaacagtgtgtggttctatctaatgtcccggtcagtaataacagttgcgacgatt 1020
OY 1433 ggggaatcctctagtgtttgacaagaagaagcaactctatgctagcttaattatccaga 1492
DB 1021 ggggaatcctctagtgtttgacaagaagaagcaactctatgctagcttaattatccaga 1080
OY 1493 ggatcgaattctacagtgctcagtcgaacaatcgaatggaattggaatcagaatatt 1552
DB 1081 ggatcgaattctacagtgctcagtcgaacaatcgaatggaattggaatcagaatatt 1140
OY 1553 gctaatgtttctcttattacgatacgtgtgcacaagtgctgcatccagattatctct 1612
DB 1141 gctaatgtttctcttattacgatacgtgtgcacaagtgctgcatccagattatctct 1200
OY 1613 cgcctcgtatcgtatctcattcgcacatgtgacacccatctgcacacccagcgggcg 1672
DB 1201 cgcctcgtatcgtatctcattcgcacatgtgacacccatctgcacacccagcgggcg 1260
OY 1673 aaccgaaacatcgcgtccatgcaagcccccgagctataataatccatgcaatgca 1732
DB 1261 aaccgaaacatcgcgtccatgcaagcccccgagctataataatccatgcaatgca 1320
OY 1733 atgcagcggttcatacatcagactcca 1760
DB 1321 atgcagcggttcatacatcagactcca 1348

RESULT 8
AAZ52236
ID AAZ52236 standard; DNA; 1096 BP.
XX
XX AAZ52236;
AC
XX
XX 18-JUL-2000 (first entry)
DT
XX
XX Maize glycine-rich protein 3 gene 5' regulatory element #3.
DE
XX
XX Maize; glycine-rich protein 3; GRP3; 5' regulatory element;
KW root specific gene expression; root abundant gene; monocotyledon;
KW pathogen resistance; pest; herbicide; growth rate; ds.
XX
OS Zea mays.
XX
XX
XX Key location/Qualifiers
FH TATA_signal 1048..1055
FT /tag= a
FT /standard_name= "TATA-box"
PD WO200015662-A1.
PD 23-MAR-2000.
XX
XX 10-SEP-1999; 99WO-EP06692.
PF 11-SEP-1998; 98EP-0117251.
PR
XX
XX (AGRI-) AGRIC TECHNOLOGY & GENETICS GMBH.
PA
XX
XX Felix G, Wulff D;
PI
XX
XX WPI; 2000-271382/23.
DR
XX
XX Nucleic acid sequence is used for cloning and expressing a root
PT specific or root abundant gene in a plant -
PT
XX
PS Claim 3; Page 50-51; 60pp; English.
XX
XX The present sequence is a 5' regulatory element of maize glycine-rich
CC protein 3 (GRP3) gene. It corresponds to positions 665-1760 of the
CC partial genomic clone zmGRP3. The present sequence is useful for cloning
CC and expressing root specific or root abundant genes in plants,
CC especially monocots which provide high expression efficiency and high
CC tissue specificity. Root preferred gene expression provides several

CC advantages to plants e.g. resistance to pathogens, pests, herbicides and
 CC adverse weather conditions, modification of growth rate and alteration of
 CC root tissue function. This sequence also provides a means of isolating
 CC related regulatory sequences of other plant species which confer root
 CC specificity to genes of interest operably linked to them.

XX Sequence 1096 BP; 288 A; 293 C; 221 G; 294 T; 0 other;

Query Match 62.3%; Score 1096; DB 21; Length 1096;
 Best Local Similarity 100.0%; Pred. No. 6.8e-311;
 Matches 1096; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 665 ctactcgctacgttagctatagcaccagagcgagatccagagggggagcaagtaggc 724
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 ctactcgctacgttagctatagcaccagagcgagatccagagggggagcaagtaggc 60
 QY 725 atggcccccccaatttgttacaccctttatacctaataatgattatgaagtgt 784
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 atggcccccccaatttgttacaccctttatacctaataatgattatgaagtgt 120
 QY 785 agaaaaatacaaatcttagatagataaaaaatctgttcaaatctctaactaatag 844
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 agaaaaatacaaatcttagatagataaaaaatctgttcaaatctctaactaatag 180
 QY 845 ctactgtcaaaattagccaaaggttagaatgatacaataattgttactataacct 904
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 ctactgtcaaaattagccaaaggttagaatgatacaataattgttactataacct 240
 QY 905 ttaataatgataatattgttagatggtcccaactagctaaacaagttacaagtatgc 964
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 ttaataatgataatattgttagatggtcccaactagctaaacaagttacaagtatgc 300
 QY 965 tagatgtttaaaatagggccaaacaacatattatagacctacgcgcgagatctctcg 1024
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 tagatgtttaaaatagggccaaacaacatattatagacctacgcgcgagatctctcg 360
 QY 1025 gttaagctctagcatctacacagctgacttaactcttccgcacacacactaaacagct 1084
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 gttaagctctagcatctacacagctgacttaactcttccgcacacacactaaacagct 420
 QY 1085 gctcgtcgccctgtttccgcacacagccatctccgataagcagctacacgactcgccc 1144
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 421 gctcgtcgccctgtttccgcacacagccatctccgataagcagctacacgactcgccc 480
 QY 1145 ctctgtagcgctcgcgtccctgttccgttttcttcgtgagctcgatcgctcgacaat 1204
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 481 ctctgtagcgctcgcgtccctgttccgttttcttcgtgagctcgatcgctcgacaat 540
 QY 1205 tagatctcgtcgtcgcaagatagactatccctcgcggtcggtgagcttgacaact 1264
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 541 tagatctcgtcgtcgcaagatagactatccctcgcggtcggtgagcttgacaact 600
 QY 1265 agaaaataatgagctctatggtgtctttttggccctcctaattttcgtctcggtccg 1324
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 601 agaaaataatgagctctatggtgtctttttggccctcctaattttcgtctcggtccg 660
 QY 1325 ccactctatactcgagagtgagcggtgagcgtatcgtaactctgaactgcaacggtat 1384
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 661 ccactctatactcgagagtgagcggtgagcgtatcgtaactctgaactgcaacggtat 720
 QY 1385 ggcgtactactaattgtcccggtgcaataatacaactgttgcgcgaagatgggaatctag 1444
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 721 ggcgtactactaattgtcccggtgcaataatacaactgttgcgcgaagatgggaatctag 780
 QY 1445 ttcttgacagaacaaggaactgctatgctagctaatatcaatccagagaatcgatttc 1504
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 781 ttcttgacagaacaaggaactgctatgctagctaatatcaatccagagaatcgatttc 840
 QY 1505 tacagtgtagctgcaacaatcgatgcaattggtgcatcagagatataatgcaattgtttt 1564
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 841 tacagtgtagctgcaacaatcgatgcaattggtgcatcagagatataatgcaattgtttt 900

QY 1565 cttaatcagatcgtgttcaacagtgatccagattcaagaattctctcgccctgatacg 1624
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 901 cttaatcagatcgtgttcaacagtgatccagattcaagaattctctcgccctgatacg 960
 QY 1625 atctcatgcacatggaacccaatctgtccaacccaacgagggggggaacccaataat 1684
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 961 atctcatgcacatggaacccaatctgtccaacccaacgagggggggaacccaataat 1020
 QY 1685 cgcgtccatgcagacccccgcagcagctataataatccatgcaatgcaatgcaagcgtc 1744
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1021 cgcgtccatgcagacccccgcagcagctataataatccatgcaatgcaatgcaagcgtc 1080
 QY 1745 atcatcatcgactcca 1760
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1081 atcatcatcgactcca 1096

RESULT 9
 AA52237 standard; DNA; 746 BP.

XX AA52237;

XX 18-JUL-2000 (first entry)

DE Maize glycine-rich protein 3 gene 5' regulatory element #4.

XX Maize; glycine-rich protein 3; GRP3; 5' regulatory element;

KW root specific gene expression; root abundant gene; monootyledon;

KW pathogen resistance; pest; herbicide; growth rate; ds.

OS Zea mays.

XX Key Location/Qualifiers

FT TATA_signal 698..705

FT FT /**tag= a

FT FT /standard_name= "TATA-box"

PN MO200015662-A1.

PN 23-MAR-2000.

PD 10-SEP-1999; 99WO-EP06692.

FE 11-SEP-1998; 98EP-0117251.

PR (AGRI-) AGRIC TECHNOLOGY & GENETICS GMBH.

XX Feix G. Wulff D.

PI WPI: 2000-271382/23.

DR Nucleic acid sequence is used for cloning and expressing a root

XX specific or root abundant gene in a plant -

PT Claim 3; Page 51; 60pp; English.

PS The present sequence is a 5' regulatory element of maize glycine-rich

XX protein 3 (GRP3) gene. It corresponds to positions 1015-1760 of the

CC partial genomic clone zmGRP3. The present sequence is useful for cloning

CC and expressing root specific or root abundant genes in plants,

CC especially monocots which provide high expression efficiency and high

CC tissue specificity. Root preferred gene expression provides several

CC advantages to plants e.g. resistance to pathogens, pests, herbicides and

CC adverse weather conditions, modification of growth rate and alteration of

CC root tissue function. This sequence also provides a means of isolating

CC related regulatory sequences of other plant species which confer root

CC specificity to genes of interest operably linked to them.

XX Sequence 746 BP; 169 A; 225 C; 160 G; 192 T; 0 other;

Query Match 42.4%; Score 746; DB 21; Length 746;

Best Local Similarity 100.0%; Pred. No. 2,66-208; Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1015	gactcttcggttcagctctctagcatctaacgtctgactcttaattctccgacacact	1074						
Db	1	gactcttcggttcagctctctagcatctaacgtctgactcttaattctccgacacact	60						
QY	1075	aaacagctctgctcgtccctcgtttccgcccacacccatcccgatccgagctcacc	1134						
Db	61	aaacagctctgctcgtccctcgtttccgcccacacccatcccgagctcacc	120						
QY	1135	gaactcgccttcgtagcgctccctcgttctgtttttctcgttagcgatccg	1194						
Db	121	gaactcgccttcgtagcgctccctcgttctgtttttctcgttagcgatccg	180						
QY	1195	ctcgacaattagatctcgtctcgcaagatagactatccctcgtcggtgagc	1254						
Db	181	ctcgacaattagatctcgtctcgcaagatagactatccctcgtcggtgagc	240						
QY	1255	ttgacgaactagaataatgatgtctcatgtgtctttttgcccctcctaatttcgt	1314						
Db	241	ttgacgaactagaataatgatgtctcatgtgtctttttgcccctcctaatttcgt	300						
QY	1315	ccctgcctccgcaactctatagctcgggagtgcggtacccgtatcgtaactgagtc	1374						
Db	301	ccctgcctccgcaactctatagctcgggagtgcggtacccgtatcgtaactgagtc	360						
QY	1375	gaacagtgatgctgactatctaatgtccgtcgcaagatatactgttgcgcaagatg	1434						
Db	361	gaacagtgatgctgactatctaatgtccgtcgcaagatatactgttgcgcaagatg	420						
QY	1435	gaactctagttttgacagaaacaaaggcaactgctatgctagctaatatccagaga	1494						
Db	421	gaactctagttttgacagaaacaaaggcaactgctatgctagctaatatccagaga	480						
QY	1495	gactgattttctaacagtgctatgctgaacaatcgatgcaattggcaatgacatatgc	1554						
Db	481	gactgattttctaacagtgctatgctgaacaatcgatgcaattggcaatgacatatgc	540						
QY	1555	taatggtttcttctatcatcgatcggtgtaacagtgcaatccagattagagttcttcg	1614						
Db	541	taatggtttcttctatcatcgatcggtgtaacagtgcaatccagattagagttcttcg	600						
QY	1615	ccctgactgactatcatcgacatgacacccatctgcacacccaaacgcggggagaa	1674						
Db	601	ccctgactgactatcatcgacatgacacccatctgcacacccaaacgcggggagaa	660						
QY	1675	ccggaacatcgctcatatgacagaccccccgagctataataatccatgcaatgcaat	1734						
Db	661	ccggaacatcgctcatatgacagaccccccgagctataataatccatgcaatgcaat	720						
QY	1735	gcagcggtcatcatcatcgactcca	1760						
Db	721	gcagcggtcatcatcatcgactcca	746						
RESULT 10									
AAZ52238									
ID	AAZ52238 standard; DNA: 554 BP.								
XX									
AC	AAZ52238:								
XX									
DT	18-JUL-2000 (first entry)								
XX									
DE	Maize glycine-rich protein 3 gene 5' regulatory element #5.								
XX									
KM	Maize; glycine-rich protein 3; GRP3; 5' regulatory element;								
KM	root specific gene expression; root abundant gene; monocotyledon;								
KM	pathogen resistance; pest; herbicide; growth rate; ds.								
XX									
OS	Zea mays.								
XX									
FH	Key Location/Qualifiers								

FT	TATA_signal	506..513
FT		/tag=a
XX		/standard_name="TATA box"
PN	WO200015662-A1.	
XX		
PD	23-MAR-2000.	
XX		
PF	10-SEP-1999;	99WO-EP06692.
XX		
PR	11-SEP-1998;	98EP-0117251.
PA	(AGRI-) AGRIC TECHNOLOGY & GENETICS GMBH.	
PI	Feix G, Wulff D:	
XX		
DR	WPI, 2000-271382/23.	
XX		
PT	Nucleic acid sequence is used for cloning and expressing a root	
PT	specific or root abundant gene in a plant	
XX		
PS	Claim 3; Page 51; 60pp; English.	
XX		

The present sequence is a 5' regulatory element of maize glycine-rich protein 3 (GRP3) gene. It corresponds to positions 1207-1760 of the partial genomic clone zmGRP3. The present sequence is useful for cloning and expressing root specific or root abundant genes in plants, especially monocots which provide high expression efficiency and high tissue specificity. Root preferred gene expression provides several advantages to plants e.g. resistance to pathogens, pests, herbicides and adverse weather conditions, modification of growth rate and alteration of root tissue function. This sequence also provides a means of isolating related regulatory sequences of other plant species which confer root specificity to genes of interest operably linked to them.

Sequence 554 BP; 138 A; 149 C; 127 G; 140 T; 0 other;

Query Match									
Best Local Similarity 31.5%; Score 554; DB 21; Length 554;									
Matches 554; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1207	gactcgcctcgctcgcaacgatatgactatccctcgcgtgctggtgacttgcgaactag	1266						
Db	1	gactcgcctcgctcgcaacgatatgactatccctcgcgtgctggtgacttgcgaactag	60						
QY	1267	aaaattatgctctcaatgctgtctttttggccctcctaattttcgtcctgcgcgc	1326						
Db	61	aaaattatgctctcaatgctgtctttttggccctcctaattttcgtcctgcgcgc	120						
QY	1327	actctatagctcggggaatgctgggttaccgtatcgtaacttgaagtgcgaacagtatgg	1386						
Db	121	actctatagctcggggaatgctgggttaccgtatcgtaacttgaagtgcgaacagtatgg	180						
QY	1387	cgtaactatcaatgctcccgtaacgaatacaacttctgcagcagatgggaactctagtt	1446						
Db	181	cgtaactatcaatgctcccgtaacgaatacaacttctgcagcagatgggaactctagtt	240						
QY	1447	ttgacagaacccaaggaactgctatgctagctaatatccacagagatcgatttcta	1506						
Db	241	ttgacagaacccaaggaactgctatgctagctaatatccacagagatcgatttcta	300						
QY	1507	cagtgctagctcgcaacatcgatgcaatttggcatcagaacgatatatgctaatgtttct	1566						
Db	301	cagtgctagctcgcaacatcgatgcaatttggcatcagaacgatatatgctaatgtttct	360						
QY	1567	ttatcgatacgttggtcaacaatgctgcatccagattacaagaattctcgcctgtagat	1626						
Db	361	ttatcgatacgttggtcaacaatgctgcatccagattacaagaattctcgcctgtagat	420						
QY	1627	ctcatcgacatggaacccaatctgcacaacccaacgctggggcggaacccaacatcg	1686						
Db	421	ctcatcgacatggaacccaatctgcacaacccaacgctggggcggaacccaacatcg	480						

CC root tissue function. This sequence also provides a means of isolating
CC related regulatory sequences of other plant species which confer root
CC specificity to genes of interest operably linked to them.

XX Sequence 190 BP; 50 A; 65 C; 42 G; 33 T; 0 other;

Query Match 10.8%; Score 190; DB 21; Length 190;
Best Local Similarity 100.0%; Pred. No. 1,5e-45;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1571 cgatcgtggtcaacagtcgagtcacagattctctcgcctgcatcgatctca 1630
|||
Db 1 cgatcgtggtcaacagtcgagtcacagattctctcgcctgcatcgatctca 60
OY 1631 tcgcacatggagaccacatctgccaacacgagggcgagggaacccgaacatgcgtc 1690
|||
Db 61 tcgcacatggagaccacatctgccaacacgagggcgagggaacccgaacatgcgtc 120
OY 1691 catgacagacccccacgagtcataataccatgcaatgcaatgagcggtcatc 1750
|||
Db 121 catgacagacccccacgagtcataataccatgcaatgcaatgagcggtcatc 180

OY 1751 atcgactcca 1760
|||
Db 181 atcgactcca 190

RESULT 13
AAS46518/C
ID AAS46518 standard; DNA; 5026 BP.

XX AAS46518;
XX
XX 18-DEC-2001 (first entry)

DT Tumour suppressor gene derived chemically modified sequence #240.

XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.

XX Homo sapiens.

XX WO200168912-A2.

XX 20-SEP-2001.

XX 15-MAR-2001; 2001WO-EP02955.

XX 15-MAR-2000; 2000DE-1013847.

XX 06-APR-2000; 2000DE-1019058.

XX 07-APR-2000; 2000DE-1019173.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI: 2001-602752/68.

XX Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer

XX Claim 1; SEQ ID No 240; 27pp; English.

XX The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since

CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (SS) and sequences complementary to (SS). The nucleic acid may be a
CC peptide nucleic acid-oligonucleotide (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantages to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes. Sequences with even numbered Seq ID numbers are the
CC complementary sequence of the corresponding odd numbered sequence (e.g.
CC ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
CC is missing).
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 5026 BP; 1467 A; 121 C; 1122 G; 2316 T; 0 other;

Query Match 2.8%; Score 48.8; DB 22; Length 5026;
Best Local Similarity 47.8%; Pred. No. 0.0024;
Matches 172; Conservative 0; Mismatches 187; Indels 1; Gaps 1;

OY 647 ttattaaactgagctactactgctagtcgtagcagccagcgagtcaga 706
|||
Db 2597 TTTCTTTAATAATTACCTATCTCATTAATACATACAAAACATCTATACCA 2538

OY 707 gggggggcaagtaggcacatggcccccctcaattgtgacaccccttatacctaaga 766
|||

Db 2537 ATTTTCAAAATAATTAACCTCCACCAACAATAATTAATAATTAACATATAC 2478

OY 767 ttattagatgaagtgtgaaaaaataaatttagatgatatataaattctgttct 825
|||

Db 2477 TTATCAATACCTTAATTAATAAATCCCTTTTAATTAACCATCTTAATTAATTAAT 2418

OY 826 aaatcctaaactaagctagctgctaaattagccaaagtgtagagatgagcaac 885
|||

Db 2417 ATCTCATTAATTTCAATATACCAATATTAATAACTCTTTTCATATTAACCCATTA 2358

OY 886 taattgttagctatacccttaataagatattgttagatggcccaactagctaaa 945
|||

Db 2357 CGATCTTAATAATTCAACCTTCAAAAACCATATATATTAATTAATTAACCAAAA 2298

OY 946 accagttacagttgctctagatgttttagaacaatggcctaaacacatattagagcc 1005
|||

Db 2297 ACTAATATATACACATTAATATATTAACCAATTAACCAACATTAACCACTAA 2238

RESULT 14
ABL33775/C
ID ABL33775 standard; DNA; 7503 BP.

XX ABL33775;
XX
XX 26-MAR-2002 (first entry)

DT Human immune system associated gene SEQ ID NO: 1748.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytosinatic; noctropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antiporiatic;

KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.

XX Homo sapiens.
OS
XX
XX WO200200928-A2.
PN
XX
XX 03-JAN-2002.
PD
XX
XX 02-JUL-2001; 2001WO-EP07537.
PF
XX
XX 30-JUN-2000; 2000DE-1032529.
PR
XX 01-SEP-2000; 2000DE-1043826.
PA
XX (EPiG-) EPIGENOMICS AG.
PI
XX Olek A, Piepenbrock C, Berlin K;
DR
XX WPI; 2002-130909/17.
PT
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
PS
XX Claim 1; SEQ ID NO 1748; 32pp + Sequence Listing; German.
CC
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 7503 BP; 2091 A; 174 C; 1582 G; 3656 T; 0 other;

Query Match 2.8%; Score 48.4; DB 24; Length 7503;
Best Local Similarity 49.2%; Pred. No. 0.004;
Matches 127; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
QY 745 tacaaccttatacctaagattattagtagtaagtagtgaagaaatacaatttag 804
DB 6349 TAAACATTTAAATTAATTAATCTTTAAAAAATAACAAAAAATTCAT 6290
QY 805 atgataaaatcttcttaaatcttaactaagtagtagtcttaaatagca 864
DB 6289 TTACAAATTAATTTAAAAAACTACTTACCAATTAATTAATTAATTAACA 6230
QY 865 aaaggttagaagatgatacctaattgttagctatacccttaataagtaattgt 924
DB 6229 AAAACACTTATTCTTAATCTACTATACATTAATTAATTAACAAATTAATCA 6170
QY 925 agatgagcccaactagctaaacagcttaacagcttagcttagttagaacaatgccc 984
DB 6169 AATTCATCAACAAACACCAAAATTAATTTATATCTACCAATTAATTTAAAAAC 6110
QY 985 taaacaacacataatga 1002
DB 6109 TAAAAAAACCAATTATTA 6092

RESULT 15
ABL33796/c
ID ABL33796 standard; DNA: 5416 BP.
AC ABL33796;
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Human immune system associated gene SEQ ID NO: 1769.
DE
XX
XX Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antihaemic; cyostatic; nootropic;

KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antihemetic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200200928-A2.
PN
XX
XX 03-JAN-2002.
PD
XX
XX 02-JUL-2001; 2001WO-EP07537.
PF
XX
XX 30-JUN-2000; 2000DE-1032529.
PR
XX 01-SEP-2000; 2000DE-1043826.
PA
XX (EPiG-) EPIGENOMICS AG.
PI
XX Olek A, Piepenbrock C, Berlin K;
DR
XX WPI; 2002-130909/17.
PT
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
PS
XX Claim 1; SEQ ID NO 1769; 32pp + Sequence Listing; German.
CC
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 5416 BP; 1476 A; 96 C; 1134 G; 2710 T; 0 other;

Query Match 2.7%; Score 47.2; DB 24; Length 5416;
Best Local Similarity 50.3%; Pred. No. 0.0074;
Matches 144; Conservative 0; Mismatches 138; Indels 4; Gaps 1;
QY 715 aaagttagccatgccccccctcaatttctgaaccccttatccctaagtatttag 774
DB 3899 AATACGCCCAATCACATCATTAATCTTTTAAATTCATGTAATCCAAA 3840
QY 775 tattagtgtagaaaatacaaatltagatagataaaaacttgctcaactccta 834
DB 3839 TATTAATTTACCAACGTATAAATATCCCATATATTAATATATATATATATTC 3780
QY 835 aaactaagtagtctgaataaagccaagaagtttagaagtagtaactaatgtta 894
DB 3779 AATA----ATCTATTCCTTTAAATCATTAATTAATTAATTAATCAAAATATCTTA 3724
QY 895 gctatacccttaataagatattatgttagatagagcccaaccctagccaacacgtaaa 954
DB 3723 TTACTAAATTTATATATTTCTAATTTATTTAAAAATCTTTAAAAAATCACTACTTA 3664
QY 955 cagttagcttagatgtttagaacaatgagcccaacacacacatatat 1000
DB 3663 TATAAACCCTACAAATTTTATTAATCTTAATAATTAATTAATTAATTAATTAAT 3618

Search completed: July 30, 2002, 18:07:15
Job time: 4385 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2002, 16:50:25 ; Search time 56.36 Seconds
(without alignments)
7670.599 Million cell updates/sec

Title: US-09-786-835-1

Perfect score: 1760

Sequence: 1 ctcgagatgttgatagaag.....ggtcacatcatcactactcca 1760

Scoring table: IDENTITY_NUC

Searched: Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCrUS.COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	62.2	3.5	7218	1 US-08-232-463-14	Sequence 14, Appl
2	41.4	2.4	1183	2 US-08-731-722-8	Sequence 8, Appl
3	39.6	2.2	1183	2 US-08-731-722-8	Sequence 8, Appl
4	39.6	2.2	2110	4 US-09-419-459-1	Sequence 1, Appl
5	37.4	2.1	7218	1 US-08-232-463-14	Sequence 14, Appl
6	37.2	2.1	625	4 US-08-988-416-919	Sequence 919, App
7	37.2	2.1	713	4 US-08-998-416-956	Sequence 956, App
8	36.8	2.1	8920	2 US-08-446-855A-1	Sequence 1, Appl
9	36.8	2.1	8920	4 US-09-150-741-1	Sequence 1, Appl
10	36	2.0	660	1 US-07-991-867B-32	Sequence 32, Appl
11	36	2.0	660	1 US-08-107-755A-32	Sequence 32, Appl
12	36	2.0	660	2 US-08-544-332-32	Sequence 32, Appl
13	36	2.0	1511	1 US-07-991-867B-8	Sequence 8, Appl
14	36	2.0	1511	1 US-08-107-755A-8	Sequence 8, Appl
15	36	2.0	1511	2 US-08-544-332-8	Sequence 8, Appl
16	36	2.0	4810	3 US-08-852-629-11	Sequence 11, Appl
17	36	2.0	4838	3 US-08-852-629-15	Sequence 15, Appl
18	36	2.0	80246	4 US-09-078-294-4	Sequence 4, Appl
19	35.8	2.0	777	4 US-09-007-119-12	Sequence 12, Appl
20	35.8	2.0	8700	2 US-08-392-625-16	Sequence 16, Appl
21	35.8	2.0	8700	2 US-08-466-961A-16	Sequence 16, Appl
22	35.8	2.0	8700	2 US-08-645-193B-18	Sequence 18, Appl
23	35.8	2.0	10564	1 US-08-206-176-5	Sequence 5, Appl
24	35.6	2.0	2273	3 US-08-714-918-40	Sequence 40, Appl
25	35.6	2.0	2273	4 US-09-265-315-40	Sequence 40, Appl
26	35.6	2.0	2273	4 US-09-265-315-40	Sequence 40, Appl
27	35.6	2.0	2273	4 US-09-266-417-40	Sequence 40, Appl

ALIGNMENTS

28	35.6	2.0	4291	2 US-08-417-210A-81	Sequence 81, Appl
29	35.6	2.0	12730	4 US-09-004-838-91	Sequence 91, Appl
30	35.4	2.0	1850	3 US-08-617-860B-32	Sequence 32, Appl
31	35.4	2.0	4098	2 US-08-605-106-4	Sequence 4, Appl
32	35.2	2.0	731	4 US-08-451-405A-2	Sequence 2, Appl
33	35.2	2.0	1103	4 US-09-122-400B-20	Sequence 20, Appl
34	35.2	2.0	1167	1 US-08-592-658-1	Sequence 1, Appl
35	34.8	2.0	1395	1 US-07-991-867B-25	Sequence 25, Appl
36	34.8	2.0	1395	1 US-08-107-755A-25	Sequence 25, Appl
37	34.8	2.0	1395	2 US-08-544-332-25	Sequence 25, Appl
38	34.8	2.0	2803	3 US-08-949-588-1	Sequence 1, Appl
39	34.8	2.0	6768	1 US-08-107-755A-1	Sequence 1, Appl
40	34.8	2.0	8457	1 US-07-991-867B-1	Sequence 1, Appl
41	34.8	2.0	8457	2 US-08-544-332-1	Sequence 1, Appl
42	34.6	2.0	8700	2 US-08-392-625-16	Sequence 16, Appl
43	34.6	2.0	8700	2 US-08-466-961A-16	Sequence 16, Appl
44	34.6	2.0	8700	2 US-08-645-193B-18	Sequence 18, Appl
45	34.4	2.0	80595	4 US-09-078-294-3	Sequence 3, Appl

RESULT 1

US-08-232-463-14

Sequence 14, Application US/08232463

Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.

APPLICANT: SCHEFFLINGER, F.

APPLICANT: FALKNER, F. G.

TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/232,463

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/935,313

FILING DATE:

APPLICATION NUMBER: EP 91 114 300.6

FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 30472/114 IMMU

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

CLONE: pTZ9pC-Fls

US-08-232-463-14


```

RESULT 4
US-09-419-459-1
: Sequence 1, Application US/09419459
: Patent No. 622028
: GENERAL INFORMATION:
: APPLICANT: Liu, Jin-Hao
: APPLICANT: Cheng, Kuo-Joan
: APPLICANT: Tsai, Cheng-Fang
: APPLICANT: Chang, Chia-Chih
: TITLE OF INVENTION: CELLULASE ENZYMES
: FILE REFERENCE: 08919/037001
: CURRENT APPLICATION NUMBER: US/09/419,459
: CURRENT FILING DATE: 1999-10-15
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 2110
: TYPE: DNA
: ORGANISM: Pitomyces rhiziniflata
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (3)...(1749)
: US-09-419-459-1

```

```

/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22313-0299
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/232,463
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/935,313
/ FILING DATE:
/ APPLICATION NUMBER: EP 91 114 300. 6
/ FILING DATE: 26-AUG-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 30472/114 IMMU
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703)836-9300
/ TELEFAX: (703)683-4109
/ TELEX: 899149
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7218 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ CLONE: PT9gpt-F15
US-08-232-463-14

Query Match      2.1%; Score 37.4; DB 1; Length 7218;
Best Local Similarity 2.4%; Pred. No. 0.74;
Matches 8; Conservative 188; Mismatches 139; Indels 0; Gaps 0.

OY  4 gagatgttgagagaagaagaaatcgcgcgaacgcagagctcctcaagcgcgcgaacgcgc 63
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db  1386 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1327

OY  64 atcctgcgcagtgagtgctgcgtatgctgcgaagagagctcgatgtgtcgtcatgc 123
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db  1326 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1267

OY  124 gagcttcgatttgcctccatgaccccgagatggtttgcgcgcgcgcgtgtaatgacgaga 183
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db  1266 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1207

OY  184 cgcgcgaccgcgaagaccggtgtctcgcataccatgtttagacatgacgcagcagacaacg 243
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db  1206 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1147

OY  244 cgaacacatccgagagacatagttgggggagagacagatagcttggagagaagaagaac 303
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db  1146 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1087

OY  304 aatagtgcacgctagatctcctcctgtgtccttc 338
    ::::: ::::: ::::: || || || || ||
Db  1086 RRRRRRRRRRRRRRRRRRRRRRATCGCAAGCTCCCTC 1052

RESULT 6
US-08-998-416-919
/ Sequence 919, Application US/08998416
/ Patent No. 6239264
/ GENERAL INFORMATION:
/ APPLICANT: Philippsen, Peter
/ APPLICANT: Pohlmann, Rainer
/ APPLICANT: Steiner, Sabine
```

```

APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reibischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSTYP11
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264rtis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Melgs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/R/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 919:
SEQUENCE CHARACTERISTICS:
LENGTH: 625 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1572RP
US-08-998-416-919

Query Match 2.1%; Score 37.2; DB 4; Length 625;
Best Local Similarity 61.2%; Fred. No. 0.24;
Matches 60; Conservative 0; Mismatches 38; Indels 0; Gaps 0.

Db 762 aatggttatttggattgaatgtagaagaacacaaattttagatagataaaactctg 821
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 194 AATATATTTATTTTATTCTTAATAATTAATAATTAATTAATTAATTAATTAATATATTC 253

Oy 822 ttctaactcctaaactaatagctagttgctcaaat 859
      || || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 254 ATTAATACTTTAATATTAATTAATTAATTAATTAATTAATTAATTAATTC 291

RESULT 7
US-08-998-416-956
Sequence 956, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reibischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSTYP11
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264rtis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 956:
SEQUENCE CHARACTERISTICS:
LENGTH: 713 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1590UP
US-08-998-416-956

Query Match 2.1%; Score 37.2; DB 4; Length 713;
Best Local Similarity 61.2%; Pred. No. 0.26;
Matches 60; Conservative 0; Mismatches 38; Indels 0; Gaps 0.

QY 762 aaatgattatcttgtaagtgtagaanaacaaattttagatagataaanaactctg 821
|| ||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 194 AATATTTTATTTTATTTTACTAAATAAATAAATAATTAATTAATTAATTAATTAATTC 253

QY 822 ttctaaatcctaaactaatagctagttgcctcaaat 859
|| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 254 ATTAAATACCTTAATATTAATTAATTAATTAATTAATTAATTC 291

RESULT 8
; Sequence 1, Application US/08446655A
; Patent No. 5849573
; GENERAL INFORMATION:
; APPLICANT: Stewart, Thomas S
; APPLICANT: Flores, Maria V
; APPLICANT: O'Sullivan, William J
; TITLE OF INVENTION: Nucleotide sequence encoding carboxymyl
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye PC
; STREET: 1100 No. 5849573th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
;

```

Query Match	2.1%	Score 36.8;	DB 4;	length 8920;
Best Local Similarity	48.1%;	Pred. No. 1,2;		
Matches 104; Conservative	0;	Mismatches 112;	Indels 0;	Gaps 0;

Query Match	2.0%;	Score 36;	DB 1;	Length 660;
Best Local Similarity	62.0%;	Pred. No. 0 57;		
Matches	57;	Conservative	0	Mismatches 25;
			100%	0
				0


```

15 RESULT 15
US-08-544-332-8
: Sequence 8, Application US/08544332
: Patent No. 5935777
: GENERAL INFORMATION:
: APPLICANT: Moyer, Richard W.
: APPLICANT: Hall, Richard L.
: APPLICANT: Grudl, Michael E.
: TITLE OF INVENTION: No. 5935777e1 Entomopoxvirus Expression System
: NUMBER OF SEQUENCES: 77
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Gerard H. Bencen
: STREET: 2421 N.W. 41st Street, Suite A-1
: CITY: Gainesville
: STATE: FL
: COUNTRY: USA
: ZIP: 32606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/544,332
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/991,867
: FILING DATE: 07-DEC-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/107,755
: FILING DATE: 19-AUG-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO 92/14818
: FILING DATE: 12-FEB-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/827,685
: FILING DATE: 30-JAN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/657,584
: FILING DATE: 19-FEB-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Bencen, Gerard H.
: REGISTRATION NUMBER: 35,746
: REFERENCE/DOCKET NUMBER: 0F114.C4
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 904-375-8100
: TELEFAX: 904-372-5800
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1511 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: Amsacta moorei entemopoxvirus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: complement (18..218)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: complement (234..782)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 852..1511
:
US-08-544-332-8

```

Oy	782	tgtgtaaaatccaatttgaataglaaaaaccttgctcaaatotcnaaacca	841
Dd	1340	TATGAAAAATTATTTATTTTAAAAAAATTAGATATATCTTCATGTATAAAAAAATAA	1399
Oy	842	tagctagtctgtaaattagccaaaggltta	873
Dd	1400	TATACATTTTGATTAATAATTTCCAAAAGATATAA	1451

```
Search completed: July 30, 2002, 18:03:08
Job time: 4363 sec
```

Query Match	2.0%;	Score 36;	DB 2;	Length 1511;
Best Local Similarity	62.0%;	Pred. No. 0.86;		
Matches	57;	Conservative	0;	Mismatches 35;
			Indels	0;
			Gaps	0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2002, 16:46:19 ; Search time 1679.9 Seconds
(without alignments)
14140.509 Million cell updates/sec

Title: US-09-786-835-1
Perfect score: 1760
Sequence: 1 ctgcagatgttgatagaag.....gtcaccatcatcactca 1760

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	287	16.3	583	10	BG842425 MEST30-G0
C 2	105.8	6.0	148	12	BH221366 1006100H0
C 3	72.4	4.1	444	9	AI948280 603039G10
C 4	72.4	4.1	515	10	BF727838
C 5	72.4	4.1	627	9	AW147089 707011D10
C 6	69.6	4.0	451	12	BH225777 1006128D0
C 7	61.4	3.5	465	10	BM381538 MEST536-B
C 8	53.8	3.1	997	12	CNS005TE
C 9	53.4	3.0	1190	12	CNS020N7
C 10	52.8	3.0	436	12	AZ916152 PstI_3_c2
C 11	49.4	2.8	1101	12	CNS0039G
C 12	48.8	2.8	275	12	BH229936
C 13	48.8	2.8	282	12	BH229935 1006155C0
C 14	48.4	2.8	815	12	BH597991 BOHFX52TR
C 15	48.4	2.8	1043	12	CNS0145P
C 16	48.2	2.7	770	12	AO740708 HS_5507_A
C 17	48.2	2.7	1101	12	CNS00EVL

C 18	48.2	2.7	1131	12	CNS034FO	AL27373 Tetradon
C 19	47.6	2.7	1092	12	CNS020K7	AL175696 Tetradon
C 20	47.6	2.7	1203	12	CNS015WU	AL106008 Drosophila
C 21	47.4	2.7	839	12	AO892711	AO892711 HS_4832_A
C 22	47.4	2.7	1225	12	CNS0161D	AL106171 Drosophila
C 23	47.2	2.7	1101	12	CNS0039G	AL063921 Drosophila
C 24	47	2.7	303	12	BH215811	BH215811 1006037A1
C 25	47	2.7	1101	12	CNS012BM	AL101356 Drosophila
C 26	46.8	2.7	777	12	BH602454	BH602454 BOHES41TF
C 27	46.6	2.6	560	10	B513308	B5133308 B5160011B
C 28	46.6	2.6	1101	12	CNS016L1	AL106896 Drosophila
C 29	46.6	2.6	1200	12	CNS016CO	AL106578 Drosophila
C 30	46.4	2.6	661	12	CNS020VT	AL209800 Tetradon
C 31	46	2.6	625	12	BH571050	BH571050 BOGH052TR
C 32	45.6	2.6	564	9	AI947590	AI947590 603024D11
C 33	45.6	2.6	975	12	BH179465	BH179465 014_P_10-
C 34	45.4	2.6	788	12	CNS00AK8	AL055582 Drosophila
C 35	45.4	2.6	1086	12	CNS00YXK	AL096362 Drosophila
C 36	45.2	2.6	340	12	BH413738	BH413738 1007034C0
C 37	45.2	2.6	616	9	AV672208	AV672208 AV672208
C 38	45.2	2.6	1159	12	CNS015XR	AL106041 Drosophila
C 39	45	2.6	1101	12	CNS004ZW	AL055440 Drosophila
C 40	44.8	2.5	512	10	B1451206	B1451206 Kx11h09.Y
C 41	44.8	2.5	556	12	BH446808	BH446808 BOGOT82TR
C 42	44.6	2.5	1201	12	CNS0103V	AL098485 Drosophila
C 43	44.6	2.5	1201	12	CNS016EA	AL106636 Drosophila
C 44	44.4	2.5	226	12	CNS0154S	AL104998 Drosophila
C 45	44.4	2.5	573	9	AU004580	AU004580 AU004580

ALIGNMENTS

RESULT 1
BG842425/C
LOCUS MEST30-G03.T3 ISUM4-TN Zea mays cDNA clone MEST30-G03 3', mRNA
DEFINITION
ACCESSION BG842425.2 GI:14244354
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays.
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 583)
Qiu,F., Cui,F., Guo,L., Ashlock,D.A., Wen,T.J. and Schnable,P.S.
Expressed Sequence Tags from B73 Maize Seedlings and Silks
Unpublished (2001)
On May 25, 2001 this sequence version replaced gi:14208747.
Contact: Patrick S. Schnable
Schnable Laboratory
Iowa State University
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schnable@iastate.edu

PCR Primers
FORWARD: T7-1 (AA TAC GAC TCA CTA TAG)
BACKWARD: T3 (ATT AAC CCT CAC TAA AG)
Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).
Location/Qualifiers

FEATURES
source
1..583
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="MEST30-G03"
/clone_lib="ISUM4-TN"
/tissue_type="Seedling and silk"
/lab_host="DH10B"
/note="Vector: pT7T3PAC; Site_1: EcoRI; Site_2: NotI;
ds-cDNA molecules were generated as follows. First-strand

Query Match	Similarity	4.1%	Score 72.4	DB 9	Length 627
Best Local	Similarity	72.4%	Pred. No. 4e-08		
Matches	113	Conservative	0	Mismatches	31
				Indels	12
				Gaps	1
OY	813	aaaatctgttctaatactcctaactaaactagctagttgctaaatattagccaaagttt	872		
DB	529	AAGGTCGTGTTCCAAACCTCTAAACTAATATTGTCGATTAAATTAACCTGAAGGTTT	470		
OY	873	agaatgagatca-----actaatgttactcctaactcctaataatagtttaatt	920		
DB	469	AAAGGAGTACATATATGAGACCCACTAATGTTAGCTTCACTCTCAATATGATCATAGC	410		
OY	921	tgttagatggtcccaactagctaaaccagttaca	956		
DB	409	TGTTAGCTGTCTCAACCCAGCTAAACCACTAATACA	374		
RESULT	6				
LOCUS	BH225777	451 bp	DNA	linear	GSS 08-NOV-2001
DEFINITION	1006128D01.x1 1006 - Rescuemu Grid G Zea mays genomic, DNA				
ACCESSION	BH225777				
VERSION	BH225777.1	GI:16824081			
KEYWORDS	GSS.				
SOURCE	Zea mays.				
ORGANISM	Zea mays				
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.				
AUTHORS	1 (bases 1 to 451)				
TITLE	Maize genomic sequences found using engineered Rescuemu transposon				
JOURNAL	unpublished (2001)				
COMMENT	Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Very probable ligation site of ends cut by single endonuclease. Reverse complemented post-ligation sequence from source sequence. Plate: 1006128 row: 2 Class: transposon-tagged. Location/Qualifiers 1..451 /organism="Zea mays" /cultivar="mixed background W23/Al18/B73" /db_xref="taxon:4577" /clone_lib="1006 - Rescuemu Grid G" /issue_type="leaf" /dev_stage="adult" /lab_host="DH10B" /note="Organ: leaf; Vector: Rescuemu (engineered from pBluescript backbone); Site:1: BamHI; Site:2: BglII; Rescuemu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuemu, go to the web site 'www.zmbd.lastate.edu' and follow the links for 'Rescuemu'. Grid G was grown at Stanford in 2000. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."				
FEATURES	source				
BASE COUNT	115 a	123 c	121 g	92 t	

FEATURES	source	Location/Qualifiers
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
ORGANISM		
SOURCE		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
LOCUS		
DEFINITION		
LOCUS		
DEFINITION</		

Query Match	4.08;	Score 69.6;	DB 12;	Length 451;
Best Local Similarity	68.68;	Pred. No. 1.9e-07;		
Matches 96;	Conservative 0;	Mismatches 44;	Indels 0;	Gaps 0;

```

/organism="zebra mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ME8536-B09"
/clone_lib="ISUM6"
/tissue_type="mixed"
/lab_host="DH10B"
/notes="Vector: pSL1p7 (4.43 kb). Site 1: EcoRI. Site 2:
NotI. Tissue samples were collected and partially pooled
prior to RNA extraction. First-strand cDNAs were prepared
from 21 individual pools of Oligo-dT selected mRNAs by
priming with 21 different NotI Oligo-dT tag primers
5'-AATCGAGGAAGATTCGCGCCGCGCCNNNNNTTTTTTTTTTTT-3').
Distinguishable 'bar code' tags. (N6). Were used for each
separate first-strand cDNA synthesis. Hence, these bar
code tags can be used to identify the mRNA pool from which
particular cDNA clone was derived. The 'bar code' tags

```


	SOURCE	Tetraodon nigroviridis.
ORGANISM		Tetraodon nigroviridis
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;	
	Acanthomorphia; Acanthopterygii; Percomorpha; Tetraodontiformes;	
	Tetraodontidae; Tetradon.	
REFERENCE		1 (bases 1 to 1190)
AUTHORS		Ruest-Crollius,H., Jallion,O., Dasilva,C., Fizesas,C., Fisher,C., Bonneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weisenbach,J.
TITLE		Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL		Unpublished
REFERENCE		2 (bases 1 to 1190)
AUTHORS		Ruest-Crollius,H., Jallion,O., Dasilva,C., Bonneau,L., Fisher,C., Bernot,A., Fizesas,C., Mincker,P., Brottler,F., Quetier,F., Saurin,W. and Weisenbach,J.
JOURNAL		Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
REFERENCE		Unpublished
AUTHORS		3 (bases 1 to 1190)
JOURNAL		Genoscope.
COMMENT		Direct Submission Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases This sequence is a single read and was generated as part of a large scale clone-and sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon. location/Qualifiers
FEATURES		1..1190
source		/organism="Tetraodon nigroviridis" /db_xref="taxon:99883" /clone="153P04" /clone_lib="G" /note="Genoscope sequence ID : C0AG15DH02SP1-end : puc-ori"
BASE COUNT		373 a 229 c 141 g 312 t 135 others
ORIGIN		
Query Match		3.0%; Score 53.4; DB 12; Length 1190; Best Local Similarity 39.5%; Pred. NO. 0.0054; Matches 18; Conservative 38; Mismatch 203; Indels 1; Gaps 1;
OY		766 attatttagtcatgaagcgtctagaacaaatcaaaaatttagatagtcataaaactctgtlcc 825
Dd		787 AATAAATAMWTAMTATMTTTTAAATATATMMWTAATTTTTTWTATATMTMTATT 846
OY		826 aaactctaacaataagctagcttgcttaaatgaaccaaaaggcttagaatgatgcac 885
Dd		847 WTATAAATATTAATAAATATNTTTTAAAWMATATANAATAATATATATMWTANNA 906
OY		886 taacttgctctataaccttaataataagtatlaa-ttgttagatggccaacactagctaa 944
Dd		907 TATTAATMTATATATATATAMTATATATATATATATATATATAMAATATATACATATMTATTAT 966
OY		945 aaccagcttaacagtgtagctctagaatgtttaagaacatggccttaaacacaacatatagc 1004
Dd		967 ATAAATATANAATATATATATATAMACAATATATATCAACAMACAMACAYACAMANNAT 1026
OY		1005 ctactcgcgagctctctcggttcagctcagtcagtcataccaagctcgtactcaactccccc 1064
Dd		1027 ATAATATCTCTMTGTATTCATTCMCATATMTCTCCTGTMTCTCACMCTTWCMTCTCTCC 1086
OY		1065 gaacacacataacagatctgctcagctgagccctglttcgcgcacagcacatccagatgc 1124
Dd		1087 HCCCCHCCCCCCCCCCCCCCCCAACACACCCTCTCCCCCCCCCCCCCCCCCCCCC 1146
OY		1125 gagcttaaccgaacttgccctctctagtcgctcgtgcc 1164
Dd		1147 CCCCCCCCCCCTACTCTCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 1186

RESULT	10	436 bp	DNA linear	GSS 15-MAR-2001
LOCUS	AZ916152			
DEFINITION	PstI_3-c-c-1_0 Maize PstI B73 leaf		Zea mays genomic, DNA sequence.	
ACCESSION	AZ916152			
VERSION	AZ916152.1	GI:13347428		
KEYWORDS	GSS.			
SOURCE	Zea mays.			
ORGANISM	Zea mays			
REFERENCE	Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.			
AUTHORS	1 (bases 1 to 436)			
TITLE	Missouri Maize Project--Maize Mapping Project.			
JOURNAL	PstI Zea mays B73 PstI leaf tissue library			
COMMENT	Unpublished (2001)			
CONTACT	Schroeder S			
LOCATION	Missouri Maize Project--Maize Mapping Project			
UNIVERSITY	University of Missouri			
209	Curtis Hall, Columbia, MO 65211, USA			
TEL:	573 882 8214			
FAX:	573 884 7850			
EMAIL:	sschroeder@celephais.agron.missouri.edu			
CLASS:	Shotgun.			
FEATURES	Location/Qualifiers			
SOURCE	1..436			
/organism="Zea mays"				
/cultivar="B73"				
/db_xref="taxon:4577"				
/clone_lib="Maize PstI B73 leaf"				
/tissue_type="leaf"				
/lab_host="DH5 alpha"				
/note="Organ: Leaf; Vector: pUC19; PstI digested B73 genomic sucrose gradient size fractionated fragment sizes of 0.5kb to 2kb ligated to pUC19 transformed in DH5 alpha"				
BASE COUNT	136 a 77 c 75 g 148 t			
ORIGIN				
Query Match	3.0%: Score 52.8; DB 12; Length 436;			
Best Local Similarity	67.9%; Pred. No. 0.0052;			
Matches 110: Conservative 0; Mismatches 37; Indels 15; Gaps 2;				
Qy	810 ataaaaacctgtctctaactctctaactaactatgctcagctgctgaatgacgaagaag 869			
Db	246 ATTGACCCCTTTTTCACCTCTTAACCTATGATTGTTACAGCTATTAATTCAGCATG 187			
Qy	870 tttagaatgagtcacaa-----ctaatgtctagctatacccttaataag---t 914			
Db	186 TTTGGAGACGACGACGCTAATATACCACTCAATTATGTATGTATACCTTCACATATGCTGCT 127			
Qy	915 atiaattgtatgatacgcccaacctagctaaaccagatcaac 956			
Db	126 ATTAGTTGTAGCTGCTGCACAACTCACTAAGACCACTAATA 85			
RESULT 11				
CNS00339G				
LOCUS	1101 bp	DNA linear	GSS 03-JUN-1999	
DEFINITION	Drosophila melanogaster genome survey sequence REF3 end of BAC # BAC008K10 of RCGI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.			
ACCESSION	AL0633921			
VERSION	AL0633921.1	GI:4941778		
KEYWORDS	GSS.			
SOURCE	fruit fly.			
ORGANISM	Drosophila melanogaster			
Eukaryote:	Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE	1 (bases 1 to 1101)			
AUTHORS	Genoscope.			
TITLE	Direct Submission			

